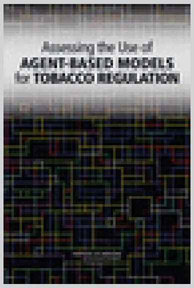

Assessing the Use of Agent-Based Models for Tobacco Regulation

ISBN
978-0-309-31722-1

280 pages
6 x 9
PAPERBACK (2015)

Robert Wallace, Amy Geller, V. Ayano Ogawa, Editors; Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation; Board on Population Health and Public Health Practice; Institute of Medicine

Assessing the Use of AGENT-BASED MODELS for TOBACCO REGULATION

Committee on the Assessment of Agent-Based Models to Inform
Tobacco Product Regulation

Board on Population Health and Public Health Practice

Robert Wallace, Amy Geller, and V. Ayano Ogawa, *Editors*

INSTITUTE OF MEDICINE
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, NW Washington, DC 20001

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This study was supported by Contract/Grant No. HHSF22301031T between the National Academy of Sciences and the Department of Health and Human Services: Food and Drug Administration. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the organizations or agencies that provided support for the project.

International Standard Book Number-13: 978-0-309-31722-1
International Standard Book Number-10: 0-309-31722-3

Additional copies of this workshop summary are available for sale from the National Academies Press, 500 Fifth Street, NW, Keck 360, Washington, DC 20001; (800) 624-6242 or (202) 334-3313; http://www.nap.edu.

For more information about the Institute of Medicine, visit the IOM home page at: **www.iom.edu.**

Printed in the United States of America

The serpent has been a symbol of long life, healing, and knowledge among almost all cultures and religions since the beginning of recorded history. The serpent adopted as a logotype by the Institute of Medicine is a relief carving from ancient Greece, now held by the Staatliche Museen in Berlin.

Suggested citation: IOM (Institute of Medicine). 2015. *Assessing the use of agent-based models for tobacco regulation.* Washington, DC: The National Academies Press.

"Knowing is not enough; we must apply.
Willing is not enough; we must do."
—Goethe

INSTITUTE OF MEDICINE
OF THE NATIONAL ACADEMIES

Advising the Nation. Improving Health.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. C. D. Mote, Jr., is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Victor J. Dzau is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. C. D. Mote, Jr., are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

COMMITTEE ON THE ASSESSMENT OF AGENT-BASED MODELS TO INFORM TOBACCO PRODUCT REGULATION

ROBERT WALLACE (*Chair*), Irene Ensminger Stecher Professor of Epidemiology and Internal Medicine, Department of Epidemiology, University of Iowa

ELIZABETH BRUCH, Assistant Professor of Sociology and Complex Systems, Assistant Research Scientist, Population Studies Center, Department of Sociology, University of Michigan

KAREN GLANZ, George A. Weiss University Professor (Penn Integrates Knowledge [PIK] Professor), Professor of Epidemiology and Nursing, Perelman School of Medicine and School of Nursing, University of Pennsylvania

CARLOS ROBERTO JAÉN, Dr. and Mrs. James L. Holly Distinguished Professor and Chairman of Family and Community Medicine, and Professor of Epidemiology and Biostatistics, University of Texas Health Science Center, San Antonio

SCOTT J. LEISCHOW, Professor of Health Sciences Research, Mayo Clinic

DOUGLAS LUKE, Professor and Director, Center for Public Health Systems Science, George Warren Brown School of Social Work, Washington University in St. Louis

CHARLES F. MANSKI, Board of Trustees Professor in Economics, Department of Economics, Northwestern University

DAVID MENDEZ, Associate Professor, Health Management and Policy, School of Public Health, University of Michigan

NANCY A. RIGOTTI, Professor of Medicine, Harvard Medical School, Director of Tobacco Research and Treatment Center, Associate Chief, General Medicine Division, Massachusetts General Hospital

DAVID SHOHAM, Associate Professor, Department of Public Health, Loyola University Chicago

JODY SINDELAR, Professor of Public Health (Policy) and Economics, Department of Health Policy and Management, School of Public Health, Yale University

MELISSA STIGLER, Associate Professor, Division of Epidemiology, Human Genetics, and Environmental Sciences, School of Public Health, University of Texas

Consultants

LAWRENCE E. BLUME, Goldwin Smith Professor of Economics, Cornell University
ROSS A. HAMMOND, Senior Fellow in Economic Studies, Brookings Institution
ALAN H. SANSTAD, Staff Scientist, Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory

IOM Staff

AMY GELLER, Study Director
V. AYANO OGAWA, Research Associate
BETTINA RITTER, Research Assistant
ANNA MARTIN, Senior Program Assistant
DORIS ROMERO, Financial Associate
ROSE MARIE MARTINEZ, Senior Board Director, Board on Population Health and Public Health Practice

Reviewers

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Research Council's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. We wish to thank the following individuals for their review of this report:

DAMON M. CENTOLA, University of Pennsylvania
SEYMOUR S. COHEN, Marine Biological Laboratory
ANA V. DIEZ ROUX, Drexel University
STEVEN N. DURLAUF, University of Wisconsin–Madison
STEPHEN EUBANK, Virginia Polytechnic Institute and State University
SANTO FORTUNATO, Aalto University
DOROTHY HATSUKAMI, University of Minnesota, Minneapolis
LEE D. HOFFER, Case Western Reserve University
ELENA G. IRWIN, Ohio State University
PAULA M. LANTZ, George Washington University
DAVID LEVY, Georgetown University
SIDNEY REDNER, Boston University

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations, nor did they see the final draft of the report before its release. The review of this report was overseen by **RONALD S. BROOKMEYER,** University of California, Los Angeles, and **ALFRED O. BERG,** University of Washington. Appointed by the National Research Council and the Institute of Medicine, they were responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

Contents

APPENDIXES

Acronyms and Abbreviations

ABM	agent-based model, agent-based modeling
Add Health	National Longitudinal Study of Adolescent Health
BRFSS	Behavioral Risk Factor Surveillance System
CDC	Centers for Disease Control and Prevention
CISNET	Cancer Intervention and Surveillance Modeling Network
CTP	Center for Tobacco Products
DPMP	Drug Policy Modelling Program
EPA	U.S. Environmental Protection Agency
FD&C Act	Food, Drug, and Cosmetics Act
FDA	U.S. Food and Drug Administration
HHS	U.S. Department of Health and Human Services
IOM	Institute of Medicine
NCI	National Cancer Institute
NIH	National Institutes of Health
NRC	National Research Council

PATH	Population Assessment of Tobacco and Health
PHEV	Plug-in hybrid electric vehicle
RCT	randomized controlled trial
SAMHSA	Substance Abuse and Mental Health Services Administration
SnapDragon	Social Network Analysis for Policy on Directed Graph Networks
SNL	Sandia National Laboratories
Tobacco Control Act	Family Smoking Prevention and Tobacco Control Act
TPSAC	Tobacco Products Scientific Advisory Committee
TUS–CPS	Tobacco Use Supplement of the Current Population Survey

Preface

Models are often used to represent complex systems. As such, they can be used to provide a framework for thinking through difficult problems, to help researchers understand factors within the complex system and their relations to specific events, to guide data collection efforts, and to identify research needs. Models are also useful for exploring policy and regulatory questions.

There is a long history of using models as one tool to guide policy making in a range of disciplines, including transportation, environmental health, energy, and health. One public health topic in which various modeling approaches have been used to address policy questions is tobacco control. Tobacco control models have been used to look at the impact of cigarette taxes, smoke-free ordinances, and smoking cessation programs. The Center for Tobacco Products (CTP) at the U.S. Food and Drug Administration (FDA) uses modeling to inform its regulatory decisions. In fact, two population models—SimSmoke and the Cancer Intervention and Surveillance Modeling Network, or CISNET, model—were used in a recent CTP-sponsored Institute of Medicine (IOM) report to study the impact of changing the minimum purchase age of tobacco products.[1] Those models, along with insights gleaned from traditional statistical and epidemiological studies, helped inform that committee's conclusions.

Although these models continue to be important in informing policy decisions by making it possible to project the impact of tobacco policies

[1] IOM (Institute of Medicine). 2015. *Public health implications of raising the minimum age of legal access to tobacco products*. Washington, DC: The National Academies Press.

and interventions at the population level (aggregate models), the majority of these models do not consider interactions at the individual level or the heterogeneity of individuals, which can be important when examining some tobacco regulatory policies. One modeling approach that CTP is exploring to address such individual interactions is the use of agent-based models (ABMs). ABMs can elucidate interactions at the individual level and thus complement population models. Existing population-level tobacco control models have provided great insight into the effects of various tobacco policies, and multiple modeling approaches are needed—and will continue to be needed—for complex problems such as understanding the impact of tobacco regulation. The results of ABMs can also be used as inputs for population models. With that in mind, it should be noted that the recommendations in this report have no relevance for the type of modeling used in the recent 2015 IOM report noted above, which focused on projections based on population models.

Through an interagency agreement, CTP commissioned Sandia National Laboratories (SNL) to develop an ABM to explore certain tobacco policies under the jurisdiction of CTP. To that end, SNL is developing a model entitled Social Network Analysis for Policy on Directed Graph Networks (SnapDragon). Thus far, the model has been developed through an iterative process and as of July 2014 was in an intermediate stage of development. CTP asked IOM to evaluate SnapDragon in its current stage of development. The Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation was created to conduct that review, which is presented in this report. CTP also asked for advice on using ABMs in the future, which the committee addresses in this report. The developers of SnapDragon provided updates on the model up until July 31, 2014. In July 2014 and January 2015, draft descriptions of the committee's technical understanding of the model[2] were sent to SNL for technical review. In the lab's January 2015 response, SNL described some updates that had been made to the model since the July 31, 2014, cutoff date and also provided information on some changes to the model that might be made in the future. However, the committee did not have an opportunity to examine any revisions to the model made after July 2014.

In addition to evaluating SnapDragon, the report offers guidance to CTP on developing and evaluating ABMs. It was beyond the scope of this report to develop specific guidance on exactly what policies or questions could be addressed using policy-relevant tobacco control ABMs. Furthermore, developing a core set of essential domains or attributes for policy-relevant

[2]The correspondence and excerpts between the committee and SNL are available in the project public access file: https://www8.nationalacademies.org/cp/ManageRequest.aspx?key=49612.

tobacco control models is a challenging venture that would vary depending on the intended purpose of the model; the time frame of this study did not allow the committee to fully explore this. In Chapter 2 the committee outlines the complex tobacco regulatory environment and the variables that need to be considered in the current regulatory environment. Each policy question to be addressed by FDA will require input from subject-matter experts as well as modeling experts to ensure that the relevant dynamics are captured in the model. This report also offers a number of issues for ABM modelers and model users to consider during model development, some important data considerations, and an evaluation framework that CTP can use in its future development of ABMs.

Acknowledgments

The committee wishes to thank the colleagues, organizations, and agencies that shared their expertise in the development of this report. Their contributions informed the committee and enhanced the quality of the report.

First, the committee thanks the Center for Tobacco Products of FDA for sponsoring the report and for bringing attention to the importance of developing models that can inform tobacco control policy. The committee expresses its deep appreciation for the SnapDragon model developers at SNL for their willingness to present the model, respond to the committee's questions, and share their experience and thoughts on the model.

The committee thanks Lawrence Blume, Ross Hammond, and Alan Sanstad for writing commissioned papers that informed the committee on the varying views concerning ABM, the practice of and pitfalls associated with ABM, and lessons learned regarding the application of ABMs. These papers were critical to the committee's deliberations.

The committee also thanks those individuals who volunteered their time and shared a wealth of information with the committee during our public meetings. Their perspectives provided valuable insights which informed the deliberations of the committee and, ultimately, the report. The meeting agendas provided in Appendix D include the names of all speakers. The committee would especially like to thank Margaret (Maggie) Eppstein of the University of Vermont and Allison Ritter and Pascal Perez of the Drug Policy Modelling Program based at the University of New South Wales in Australia for taking the time to describe the ABMs they had developed, and to answer questions from the committee and staff.

Finally, the committee expresses its gratitude to the IOM staff members who contributed to the production of this report, including study director Amy Geller, research associate V. Ayano Ogawa, research assistant Bettina Ritter, senior program assistant Anna Martin, and board director Rose Marie Martinez. Other staff members of the Board on Population Health

and Public Health Practice, including Hope Hare, Doris Romero, and Kathleen Stratton, provided important support as well. The committee is also thankful for Norman Grossblatt and Robert Pool for their editorial support as well as for the National Academies Research Center.

On a personal note, I thank the members of the IOM Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation, who volunteered their time from their very busy professional lives to participate on the committee. I deeply appreciated their dedication and commitment to help guide FDA and the nation on this critical topic.

Robert Wallace, *Chair*
Committee on the Assessment of
Agent-Based Models to Inform
Tobacco Product Regulation

Summary

Computational modeling of social processes has been used for many years in numerous disciplines for a variety of purposes, including assisting in public policy decisions. Computational models can be used to inform the regulatory process and can "range from single parameter linear relationship models to models with thousands of separate components and many billions of calculations" (NRC, 2007, p. 36). Models have been used to forecast the health effects associated with risk behaviors, including tobacco use. For example, several population dynamic models have been used to simulate the dynamics of smoking and smoking-attributed deaths in a state or nation and the effects of policies on those outcomes (HHS, 2014). Since 2009 the U.S. Food and Drug Administration (FDA) has had broad regulatory authority over tobacco products and has used models as one tool to inform its policy decision-making activities. Recently, FDA has been exploring the usefulness of a particular computational modeling approach—agent-based modeling (ABM)—to inform its policy decisions.

Thus, the FDA Center for Tobacco Products (CTP) asked the Institute of Medicine (IOM) to review an ABM developed for use by FDA; to comment on its strengths, weaknesses, and usefulness for examining various tobacco regulatory policies; and to provide recommendations on strategies to improve the model and for using ABM to inform decision making in the future. To address that request, the IOM created the Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation (see Box S-1 for the full committee statement of task).

CTP has several reasons for its interest in using ABMs to inform tobacco control policy, including their potential for exploring individual-

BOX S-1
Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation
Statement of Task

The Institute of Medicine (IOM) shall convene a committee to assess the applicability of agent-based models of tobacco use and public health as a guide to inform regulators and improve the effect of tobacco regulation policies on public health. The committee shall:

- comment on implications of using agent-based models to examine various tobacco regulatory policies
- assess the strengths and weaknesses of an agent-based model developed for the U.S. Food and Drug Administration (FDA) (to be provided by the Center for Tobacco Products [CTP]) and models currently available in the literature that have been used for similar purposes (to be identified by CTP)
- make recommendations on future directions and strategies to improve the usefulness of the model developed for or to be used by FDA, if needed

level factors that dictate tobacco use and their ability to simulate potential use patterns in an evolving market (Fultz, 2014). It is important to note that the committee formally assessed only one ABM in this report, and although lessons from the development of that model may be applied to the development of future ABMs, this report's conclusions are not indicative of the strengths or limitations of other tobacco control ABMs or of tobacco control models using other modeling approaches. This report is meant to build on the large body of work on tobacco use modeling by exploring how ABMs might be a helpful tool to add to the existing tobacco control modeling toolkit.

BACKGROUND

What Are Agent-Based Models?

An ABM is a type of computational model that is used to study complex systems by exploring how individual elements (agents) of a system behave as a function of individual characteristics and interactions with each other and the environment. Each agent interacts with other agents based on a set of rules and within an environment specified by the modeler, which leads to a set of specific outcomes, some of which may be unexpected. As ABMs can be used to explore the potential impact of policies and interventions in dynamic social and physical environments, ABMs may be a useful tool to aid in decision making by policy makers. ABMs have been used to examine other public health interventions and policies, such as for infectious diseases (Epstein et al., 2007; Lee et al., 2010) and obesity (Auchincloss et al., 2011; Orr et al., 2014; Zhang et al., 2014), but they have not been fully explored and considered in the tobacco regulatory space.

Complex Tobacco Environment

Tobacco consumption continues to be the leading cause of preventable disease and death in the United States (HHS, 2014). More than 42 million Americans, representing 18 percent of the population, currently smoke cigarettes (Agaku et al., 2014; Jamal et al., 2014). Each day more than 3,200 children under age 18 smoke their first cigarette, and more than 700 children become daily cigarette smokers (SAMHSA, 2013). Many of these youth will become addicted and suffer adverse health consequences. At the current smoking rate, 5.6 million children alive today will die prematurely from smoking-related illness (HHS, 2014).

On June 22, 2009, the Family Smoking Prevention and Tobacco Control Act (Tobacco Control Act) gave FDA the authority to regulate the manufacture, distribution, and marketing of tobacco products—specifically cigarettes, cigarette tobacco, roll-your-own tobacco, and smokeless tobacco—to protect public health and reduce tobacco use in the United States. To oversee the implementation of the law, FDA established CTP, which works to prevent tobacco product use initiation, encourage current users to quit, and reduce the overall harm caused by tobacco use. New policies and regulations must be based on available medical, scientific, and other technological evidence as appropriate for the protection of the public's health. Consequently, CTP is interested in forecasting the public health effects of potential changes in tobacco standards and other policies.

As described in Chapter 2, understanding the complicated environment in which tobacco products are used and sold is essential when attempting to model potential tobacco policies. This includes an understanding of the various tobacco products available and their addictive nature as well as tobacco use behaviors, including tobacco use initiation, progression, and cessation.

Center for Tobacco Products and Agent-Based Modeling

Through an interagency agreement between CTP and the U.S. Department of Energy, CTP commissioned Sandia National Laboratories to develop an ABM that could help CTP understand the potential impacts of a variety of tobacco control policies on population health.[1] The ABM developed by Sandia National Laboratories for CTP is titled Social Network Analysis for Policy on Directed Graph Networks (SnapDragon). The main purpose of SnapDragon is to explore the effects of various tobacco policies and interventions, such as public education campaigns, on opinion and tobacco use within social networks.

REPORT FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS

Why Use Agent-Based Models to Explore Tobacco Use?

Existing models in tobacco control have focused mostly on determining the long-term dynamics of population-level tobacco rates. These analyses have employed almost exclusively aggregate compartmental or system dynamics models, which assume a large degree of homogeneity among the population and generally do not consider interactions among members of the population. Given that smoking is largely related to social- and individual-level behaviors, it is becoming evident that these processes need to be modeled to understand the effects that a policy may have. Although analysis of survey data can help researchers identify the nature and strength of these social determinants of smoking behavior at the individual level, ABMs are needed to estimate the total population effects of those individual interactions. ABMs can account for individuals' differences and the many ways in which such individuals can influence each other to estimate the combined effect of the multiple processes that constitute tobacco use behavior. These models can also account for important feedback mechanisms that have been, for the most part, ignored by existing aggregate models.

Given the strong social component inherent to tobacco use onset, cessation, and relapse, and given the heterogeneity of those social interactions, ABMs have the potential to be an essential tool in assessing the effects of policies to control tobacco. Many of the questions FDA faces require an understanding of the underlying behavioral mechanisms involved (e.g., initiation and cessation) and would require a model of those processes before those specific questions could be explored. Within the modeling community, it is often said that models need to be motivated by a specific question to be effective. However, the processes or mechanisms underlying these policy

[1] The modeling efforts by Sandia National Laboratories under this agreement have included population health models that aim to help forecast potential long-term impacts on prevalence, morbidity, and mortality for the population in the United States as well as other types of models, including an ABM.

questions often need to be the focal point of the model before the specific question is addressed.

> Finding 2-1: The committee finds that for many tobacco control policy questions, several key underlying processes—initiation, cessation, and relapse, among others—drive overall rates of tobacco use and have a strong social interaction component. An agent-based model could be a useful tool to represent these processes.

In the case of tobacco, a useful path will be to develop models of these processes first and to then apply them to the specific policy question. This does not imply that all efforts should be put into a single model of social processes which would then be applied to many different questions. Rather, accurately representing the underlying process of initiation, cessation, and relapse is, in some cases, essential to the development of a model of tobacco use behavior.[2]

Mechanisms That Generate Feedback Between Behavior and Social Environments

Policies can backfire when they fail to account for how people change their behaviors in response to an intervention, as individuals' behaviors often depend on the behaviors of other people and on features of the social environment. A central challenge for policy makers, therefore, is to anticipate how organizations, corporations, and individuals will react to changes in incentive structures and features of the environment. However, anticipating this response can be difficult for several reasons, including limited knowledge of human behavior and the complex interactions that occur between individuals and the social environment.

Structural models,[3] to which policy makers have long looked to guide policy decision making, typically attempt to uncover behavioral relation-

[2] It is important to note that there are other features of tobacco control policy that are not directly related to initiation and cessation (e.g., tobacco companies responding to FDA regulatory changes in an attempt to undermine those changes), so the modeling decision to focus on a specific policy question versus initiation or cessation needs to be discussed early in model conceptualization.

[3] Structural models use a set of equations or rules, expressed analytically or in programming code, that describe different possible worlds. The specification of the model is dictated by theory, prior knowledge, and other inputs that determine what features of a given process to highlight and what to leave out. These assumptions, combined with data, produce a set of inferences about what will happen under a given set of conditions. This modeling approach includes (but is not limited to) macro-level simulation models, such as system dynamics, and individual-level models, such as ABMs. The appropriateness of a given modeling strategy depends on the theory brought to bear and on available empirical evidence.

ships or parameters that are invariant to specific circumstances or take those circumstances as inputs that condition behavior. Evidence across a number of policy domains suggests that if the incentives or risks associated with a given behavior are changed, people will likely behave differently. Thus, a non-superficial understanding of the incentives that drive behavior is required. To be useful for informing regulatory policy, modeling efforts must capture meaningful aspects of the social process under investigation. It is not enough to hypothesize different mechanisms and use a model to determine whether they lead to different outcomes. The model may be misspecified to the point where a "sensitivity analysis"[4] provides no information at all on the true sensitivity of model outputs to inputs (Sanstad, 2015). At higher levels of aggregation, the behavior of organizations and other coalitions are also contingent on behavioral incentives. Failure to account for those incentives may lead to unexpected and undesirable results. The goal is to identify how people's or organizations' behaviors might change under a different incentive structure.

Conclusion 3-1: The committee concludes that a deep understanding of human behavior, decision making, and incentive structures is important for agent-based models and other models that are used to understand how interdependent behaviors shape the outcomes of a given policy. Regardless of the model type, if the behavior is not plausible, the model is not likely to be informative.

Recommendation 3-1: When developing an agent-based model (or similar modeling approach), the Center for Tobacco Products should consult with subject-matter experts to identify the plausible behaviors and focal processes at work from the beginning of the model development process.

Microsimulation and Agent-Based Models

Within the domain of individual models, some scholars semantically distinguish between two types: *microsimulation* and *agent-based models*. However, both involve the same basic procedure—assigning artificial agents a behavior and using simulation to assess the aggregate implications of that behavior—and both approaches are operationalized through computer code. This is an important commonality because if microsimulation and ABMs are viewed as two distinct approaches, their two research communities will be less likely to benefit from each other's work. The committee

[4]Sensitivity analysis is "an exploration, often by numerical (rather than analytical) means, of how model outputs (particularly QOIs [quantities of interest]) are affected by changes in the inputs (parameter values, assumptions, etc.)" (NRC, 2012, p. 117).

found that from a purely technical standpoint, microsimulations and ABMs are the same modeling enterprise and that they are differentiated mainly by differences in how they tend to be deployed, including in how agent interactions are specified and how agents' environments are abstracted. Although there are no fundamental differences between ABMs and microsimulations, there are historical differences in how these models have been specified and used by their research communities.

Conclusion 3-2: Researchers who use the terms agent-based modeling and microsimulation have different approaches to model specification. However, the committee concludes that from a technical standpoint these are the same enterprise (an individual-level model implemented via computer code). The committee believes that modelers would greatly benefit from best practices and lessons learned from applications that have been performed by the two research communities to address policy questions.

High-Dimensional Models and Low-Dimensional Models

The appropriate level of model detail and empirical realism is a choice that modelers need to make. The appropriate level of model detail depends on the research question, the intended use of the model, and the data available to empirically ground the model. It is important to note, however, that at whatever level, models provide only an imperfect representation of the real world, as computational models in general are not reality mirrors, nor are they intended for this purpose. ABMs can represent anything from low-dimensional, abstract worlds where agents are defined by just one or two attributes and interact in a highly stylized environment based on simple rules to high-dimensional, highly detailed worlds where agents have many attributes, the environment contains a great deal of information, and agents engage in multiple behaviors. It may be tempting to create ABMs that pull in as much empirical data and knowledge as possible in an attempt to create a highly realistic "laboratory" to explore policy questions. However, this approach is not usually the most productive, because available data and knowledge of human behavior are almost never adequate to achieve this. ABM allows the developers to explore the importance of various mechanisms in the face of no data and to assess the potential value of collecting these data; however, this introduces an added layer of uncertainty and raises the possibility of model misspecification. Also, the model can become cumbersome and hard to manage when additional layers of detail are added, and it can be difficult to get clear analytic results. The success of a model is not determined by the level of granularity at which it represents a process; rather its success is based on how successfully it facilitates the understanding of the problem or question under study.

Conclusion 3-3: The committee concludes that low-dimensional and high-dimensional models have complementary virtues and weaknesses. A more complicated model may have greater verisimilitude, but added detail per se does not ensure realism. A low-dimensional model, while abstracting from some features of the real world, may generate forecasts that are easier to understand and interpret.

Recommendation 3-2: The Center for Tobacco Products should develop and employ both low- and high-dimensional models, using both as appropriate to shed light on policy impacts.

Making Decisions with Partial Knowledge

Models cannot predict the future with certainty. They provide only a partial representation of reality and have some level of abstraction. Models can mislead policy making if modelers present findings with greater certitude than is warranted. These challenges are pertinent to any type of model that seeks to inform policy, not just ABMs. A good model will quantify how uncertainty in model inputs translates into uncertainty in the likely outcomes of various policies and will generate a range of predictions that reflect that uncertainty (Manski, 2013; Wagner et al., 2010). The key challenge is separating what is known from what is unknown. Note that this is a very different enterprise from conducting a “parameter sweep” type sensitivity analysis, which merely provides more insight into the workings of the model itself and not into the relationship between the model and the actual world. Once analysts have generated a set of credible model outputs, they must use that information to draw a conclusion about the best course of action. The challenge for the policy maker is to evaluate candidate policy outcomes and weigh the risks and benefits. Thus, to use these models effectively to guide policy decisions, the model user needs a rule for translating these uncertain predictions into a policy decision.

Conclusion 3-4: The committee concludes that the common exercise of sensitivity analysis does not suffice to measure the uncertainty in model-based forecasts. Sensitivity analysis may provide some insight into the workings of the model itself, but it does not per se assess the potential relationship between model findings and the real world.

Recommendation 3-3: When the U.S. Food and Drug Administration uses the findings of any model, the agency should take into account the uncertainty of findings in order to evaluate policy outcomes and weigh the risks and benefits appropriately.

An Evaluation Framework for Policy-Relevant Agent-Based Models

Policy-relevant ABMs are complex, resource-intensive technical activities that are developed by large groups of people with varying areas of expertise and whose results need to be translated and communicated to various stakeholders in order to affect policy and improve health. Policy-relevant ABMs need to be built carefully using appropriate data and social science theories, rigorously tested, and clearly communicated. These requirements for ABMs are the same as for other types of computational models and simulations used to inform policy decisions. Given the amount of time, effort, and money required to build an effective policy-relevant model, it is critical to evaluate the process, its outcomes, and its overall value. Chapter 4 presents an evaluation framework for policy-relevant ABMs developed by the committee. Such an evaluation framework can help model developers improve their modeling efforts, help funders understand better how to use model results and how to guide future funding of modeling work, help policy makers understand how to translate model results into more effective policies and increase their trust in the analysis, and help modelers and scientists by suggesting new avenues for research, modeling, and data collection.

Fundamental Evaluation Categories

The evaluation framework developed by the committee has five fundamental evaluation categories that the committee believes need to be included in most ABM evaluations:

1. *Resources*: The modeling team needs access to adequate financial, infrastructure, human, and knowledge resources to successfully design, build, and test the model.
2. *Technical Best Practices*: Model implementation, testing, and validation phases need to be reviewed throughout model development.
3. *Model Suitability*: Models need to be developed in a manner that makes them suitable for their intended purpose and allows for exploration or testing of specific policy options or conditions. Some models could be developed for very narrow questions related to tobacco use, others as a broad tool to look at a larger range of tobacco policies.
4. *Communication and Translation*: Communication and translation strategies are essential during every stage of model development for enhancing the model-building process and ensuring that the model is focused on the key issues that will affect policy outcomes. Modeling requirements, descriptions, and results need to be com-

municated effectively to a variety of audiences, including agency staff, regulators, politicians, and the general public.

5. *Policy Outcomes*: Ultimately, policy-relevant models will be used to inform policy and regulatory action or to advance scientific progress.

Recommendation 4-1: The Center for Tobacco Products should adopt an evaluation framework for its modeling work, either the one presented in this report or one similar to it. Key dimensions of the evaluation framework should include considerations of resources, technical best practices, model suitability, communication and translation, and policy outcomes. The evaluation plan should be designed early in the model development process and should be carried out throughout model development.

This evaluation framework would apply to all efforts funded by CTP (internal model development, interagency agreements, contracts, and grants). The evaluation—as well as periodic peer review[5]—should come from external experts in addition to internal reviewers. If CTP chooses to adopt the framework outlined by the committee, it should be used as a guideline and not as a mechanical exercise or checklist, as different ABMs will require differing evaluation strategies based on their intended use, modeling approach, and other aspects of model development.

Review of the Social Network Analysis for Policy on Directed Graph Networks Model (SnapDragon)

As a major component of its statement of task, the committee was asked by FDA to review the FDA-commissioned ABM developed through an interagency agreement with Sandia National Laboratories, entitled Social Network Analysis for Policy on Directed Graph Networks (SnapDragon). Chapter 5 describes and analyzes SnapDragon, applying the evaluation criteria from Chapter 4 where appropriate, and discusses the model's usefulness for informing tobacco control policy. The SnapDragon model has not been published in a peer-reviewed journal, but two manuscripts on the model were undergoing peer review in two different journals during the course of this study and have since been accepted for publication (Moore

[5] See the National Research Council (2007) report on modeling for guidance on peer review of models.

et al., in press a,b).[6] The committee reviewed the model as it was specified as of July 31, 2014.[7]

To date, the SnapDragon model development team has focused primarily on the effects of multiple competing high- and low-opinion messages in social networks, illustrated through the study of education campaigns, in a single- or multiple-tobacco-product environment. The model distinguishes individuals by any number of characteristics and, in particular, according to their use of a variety of tobacco products, which allows for the investigation of the simultaneous use of different products. Currently, the model classifies individuals as either "users" or "nonusers" for each tobacco product under consideration. The user status is determined by an underlying construct termed "opinion."[8] Each individual carries an opinion about each tobacco product under consideration, which drives tobacco use behavior. SnapDragon explicitly models the time trajectory of individuals' opinions as a result of their interactions with other individuals, and the modeling choice is based on theory stemming from the field of opinion dynamics. In the model, individuals are connected to others through predefined social networks. Connected individuals can affect each other's opinions if such opinions do not differ by more than a specified tolerance range. As time goes on, individuals continuously adjust their opinions toward a weighted average of the opinions of the individuals who can influence them. The weights (or plasticity values) represent the importance given to the opinion of particular individuals. These plasticity values are not necessarily reciprocal, meaning that any two connected individuals can assign a different weight to the opinion of the other. As opinions adjust, they drive tobacco use behavior (i.e., become a user or a nonuser of a specific tobacco product). That is, nonusers whose opinions about the product increase beyond a certain level (termed the initiation threshold) become users of such tobacco product,

[6]The draft manuscripts reviewed by the committee and other supporting documents are available upon request from the project public access file: http://www8.nationalacademies.org/cp/ManageRequest.aspx?key=49612.

[7]In addition to the draft manuscripts (dated May and November 2013), the committee received more information on SnapDragon from in-person presentations by the model developers during two open information-gathering sessions as well as from written question-and-answer documents exchanged between the committee and the SnapDragon development team. In July 2014 and January 2015, draft descriptions of the committee's technical understanding of the model were sent to the model developers for technical review. In their January 2015 response, the developers noted several changes to the model that occurred after July 31, 2014, and identified additional changes they planned to make in the future. However, the review by the committee is based on the model as it existed on July 31, 2014.

[8]The modeling team defines "opinion" as an aggregate concept that captures the overall positive or negative attitude of a person toward a tobacco product. It is represented as a continuous variable, taking values between 0 and 1, with 0 standing for the most negative attitude of a person toward a tobacco product and 1 the most positive.

and users of a particular product whose opinions about such product fall below a certain level (termed the cessation threshold) become nonusers of such product. Interventions that can potentially influence individuals' behaviors, such as tobacco control efforts, are modeled as modifying either the opinions of individuals about a certain tobacco product or the opinion thresholds that delimit possible user status.[9]

Review of SnapDragon

In Chapter 5 the committee provides a detailed review of the SnapDragon model; the key findings and conclusions are described here. While SnapDragon has been designed to evaluate a wide range of tobacco products, the committee focused on how the structure of the model can accommodate known facts about smoking behavior.

SnapDragon presents a novel framework to study the impact of policy measures on smoking initiation, cessation, and prevalence by attempting to model explicitly the processes of initiation and cessation as driven by social interactions. Instead of relying on externally supplied inputs for initiation and cessation rates that were determined outside the model, the model tries to derive these figures endogenously, by proposing a hypothesis about how these processes are generated. Specifically, SnapDragon attempts to explain the dynamics of tobacco use (i.e., how the system changes over time) as a result of a convergence of opinions about specific tobacco products through the interaction among individuals in the population, guided by a formulation from the field of opinion dynamics (see Chapter 5).

However, the committee found that several elements in SnapDragon's formulation either do not conform to existing knowledge about tobacco use or defy face validity. First, the model does not consider a feedback mechanism from behavior to opinion. It is almost certain that the experience of using a particular tobacco product would influence the user's opinion about such product. As SnapDragon only considers that behaviors are modified through opinions, this suggests that the model is missing an important feedback mechanism from behavior to opinion. Second, it is very unlikely that opinions about tobacco products are transmitted independently of individuals' behavior toward such products, as SnapDragon stipulates. This formulation could lead to highly unrealistic scenarios, as explained in Chapter 5.

It is conceivable that imitation of smoking behavior could play a role in tobacco adoption. However, in SnapDragon the imitation component happens indirectly—by individuals sharing their opinions about a product—rather than directly through behavior. As opinions are not

[9] A detailed description of SnapDragon is available in Chapter 5.

influenced by behaviors in the model, the observed number of tobacco users in the population will not affect the rates at which new adopters are generated. Similarly, observed quitting behavior cannot be imitated directly in the model.

Third, the rationale behind the modeling choice of making interacting opinions converge to a weighted average is not clear. This modeling choice, when applied to smoking behavior, can lead to inconsistencies with observed facts. It is likely that other mechanisms, not reflected in SnapDragon, play an important role in modifying smoking behavior throughout an individual's lifespan. Furthermore, SnapDragon does not identify former smokers. As such, the model in its current form can track product prevalence but cannot accurately determine health risks,[10] as a significant proportion of tobacco-related morbidity and mortality falls on former users of combustible products (HHS, 2014). These and other limitations of the model—including other aspects of tobacco use behavior, equilibrium patterns, and the use of data in the model—are outlined in Chapter 5.

The developers of SnapDragon have suggested that it could be used for a number of tobacco control policy applications, but the underlying assumptions of the model (as discussed in Chapter 5) suggest that this is unlikely. The committee statement of task calls for recommendations for improvement of SnapDragon, if needed, and although some changes could be made to address some of the weaknesses identified in this report, doing so would lead to the creation of a new model. SnapDragon does not encompass essential facts from the tobacco research literature, and many of its assumptions lack face validity. In addition, the data required to inform the parameters in SnapDragon have not yet been identified, and the model has not yet reached the stage of model validation for broad application to tobacco control policy. While SnapDragon is a very flexible model that can be manipulated in various ways to reproduce certain observed facts about tobacco use behavior, it currently lacks sufficient modeling structure to be informative for policy. Therefore, the committee has not included recommendations for improvement. Key findings and conclusions regarding SnapDragon are below:

> *Conclusion 5-1: As SnapDragon presumes that opinions may modify behavior but behavior does not modify opinion, the committee concludes that the model is missing an important feedback mechanism from behavior to opinion.*

[10] Although determining health risks was not listed as one of the purposes of SnapDragon, if CTP plans to use SnapDragon as a stand-alone model, this is a limitation. If CTP plans to use SnapDragon only to inform population models, this is not a limitation of the model.

Finding 5-1: The committee finds that the representation of behavior in SnapDragon does not align with what is currently known about tobacco use and dependence.

Conclusion 5-2: The committee concludes that the modeling decision of making interacting opinions about tobacco converge to a weighted average is not supported by evidence and is unlikely to be an accurate representation of tobacco use behavior.

Finding 5-2: Whereas some other models based on opinion dynamics have been able to replicate the equilibrium patterns of socially driven processes, the committee has not found applications in which the specific time path to equilibrium has been empirically validated.

Finding 5-3: The committee finds that there has been no assessment of SnapDragon's ability to accurately predict initiation, prevalence, or cessation.

Conclusion 5-3: The committee concludes that a realistic parameterization of SnapDragon would be hard to achieve, so it is unlikely that the model will be able to generate credible assessments of policies.

Recommendation 5-1: SnapDragon should not be pursued by the Center for Tobacco Products as an aid for regulatory decision making.

Data Collection and Model Development at the Center for Tobacco Products

Chapter 6 provides a high-level overview of existing tobacco use data sources, identifies data gaps, and makes recommendations for the future implementation of ABM at CTP. Various types of existing data sources related to tobacco use can be used to inform and strengthen ABMs, but these sources do not contain all relevant agent attributes, behaviors, and social and spatial interactions related to tobacco use. One approach to access such data would be to try to maximize the use of available administrative data from states and regions, but except in unusual circumstances, this information is unlikely to contain many of the behaviors and interactions wanted. Another approach would be to combine data from various sources, such as large-area administrative information and small-area detailed surveys. However, using such combinations would require considerable care. A longer-term approach would be to try to anticipate critical data needs and fund or otherwise encourage the collection of data that best suit ABMs or other modeling approaches. Similarly, encouraging the standardization of data collection items and methods might improve model quality. Even for administrative data that are "routinely" collected, such as tobacco marketing or sales information or population smoking prevalence estimates, it could be possible to evaluate those data periodically for validity and consistency. It may also be possible to substitute existing or newly developed biomarkers of certain smoking behaviors for other forms of data collection, and, in selected instances, information from other countries with similar populations may be of value.

The committee also discusses the importance of collecting data that inform agent interactions, either with other agents or with the agent's environment, which are a key element in ABMs. Such interactions are difficult or impossible to capture empirically, but alternative data collection methodologies, including qualitative methods and experiential or situational sampling, could help overcome this challenge. Because ABMs and other individual-level modeling techniques are promising tools to further the understanding of tobacco use behavior, it is worthwhile to collect such data. As a major funder and user of tobacco data (including for the modeling of tobacco use), CTP can help shape the tobacco data environment in the future.

Conclusion 6-1: The committee concludes that agent-based models designed to inform policy decisions require data on the underlying mechanisms governing behavior and on agent-to-agent and agent-to-environment interactions. Currently, these data are not commonly collected.

Recommendation 6-1: The Center for Tobacco Products should identify and help develop data sources relevant to the questions it is trying to address using agent-based models and other modeling approaches.

Data already being collected (either by CTP or other sources) could be incorporated into the modeling process. CTP could consider coordinating with other activities, such as the Tobacco Centers of Regulatory Science, to gather this data.

To ensure that the processes of collecting the necessary data and identifying agent attributes based on those data are done successfully, it is crucial to address implementation issues. Funders for policy-relevant models require access to expertise if they are to issue effective funding opportunity announcements or contracts; to determine which modeling approaches are appropriate for the question at hand; to work with sponsored modeling teams throughout model development; to evaluate model inputs, processes, and outputs; and to appropriately interpret model results and translate them for decision makers.

FDA is regularly confronted with uncertainty within the complex tobacco environment. Because of this, the agency will continue to need models that represent potential tobacco policies in order to organize data, elucidate uncertainties, and forecast future scenarios. Because the use of models at CTP has the potential to affect regulatory decision making, it is essential that the development of these models be overseen by individuals who have the expertise and experience needed to maximize the benefit and reliability of the models.

Recommendation 6-2: The Center for Tobacco Products (CTP) should ensure that it has staff with, or access to, the necessary expertise to inform CTP's research, contracting, and evaluation efforts and to translate model results for various stakeholders.

Although individual models are a useful tool for informing policy decisions, having a range of modeling techniques will offer a fuller picture of the policy questions confronted by CTP—for example, by creating various models to approach the same question or process (for example, multiple ABMs or ABMs and aggregate models). The documentation of model inputs, activities, and outputs by the model developers (as discussed in Chapter 4) and a comparison of results with a rigorous discussion by the developers on why the results differ—or do not differ—will create a richer understanding of the models and the model results and will help to address model uncertainty. Doing so will help to increase policy makers' confidence in the model results or identify where assumptions need to be modified or where further data is needed.

Recommendation 6-3: The U.S. Food and Drug Administration should develop a range of models using various approaches. This would include agent-based models as well as other modeling approaches.

It is important to note that the range of models FDA could use includes not only those that FDA commissions or develops but also those that others have already developed or will develop to help guide tobacco control policy.

CONCLUSION

Although simulation modeling has been used for many years in tobacco control, CTP is still early in its efforts to use ABM to explore tobacco control policy and regulation. This report illustrates many of the challenging and technical aspects surrounding ABMs. However, the committee believes that ABMs are a useful tool and that they could add to the understanding of tobacco use initiation, cessation, and relapse processes. While the model developed for FDA (see Chapter 5) does not accurately represent many of the important characteristics of tobacco use, there is much that can be learned from its development that could be applied to future models of tobacco use. There are some barriers to overcome, such as the collection of data to inform the development of ABMs and understanding the empirical and theoretical challenges of specifying model inputs and appropriately interpreting model outputs (see Chapter 3). A strong evaluation framework (see Chapter 4) is needed to track model development. As discussed in Chapters 3 and 4, it will be important to consult an interdisciplinary modeling team, and subject-matter experts will need to be consulted at the earliest stage of model conceptualization and throughout the model development process to ensure that the model is grounded in the current state of tobacco science (that is, evidence-based research related to tobacco in the fields of epidemiology, social and behavioral sciences, biology, chemistry, and others), while carefully considering individual behavior. If the principles discussed in this report are followed, the role of ABMs for informing tobacco regulation will be greatly strengthened.

REFERENCES

Agaku, I. T., B. A. King, C. G. Husten, R. Bunnell, B. K. Ambrose, S. S. Hu, E. Holder-Hayes, H. R. Day, and Centers for Disease Control and Prevention. 2014. Tobacco product use among adults—United States, 2012–2013. *Morbidity and Mortality Weekly Report* 63(25):542–547.

Auchincloss, A. H., R. L. Riolo, D. G. Brown, J. Cook, and A. V. Diez Roux. 2011. An agent-based model of income inequalities in diet in the context of residential segregation. *American Journal of Preventive Medicine* 40(3):303–311.

Epstein, J. M., D. M. Goedecke, F. Yu, R. J. Morris, D. K. Wagener, and G. V. Bobashev. 2007. Controlling pandemic flu: The value of international air travel restrictions. *PLoS ONE* 2(5):e401.

Fultz, S. L. 2014. *Committee on assessment of agent-based models to inform tobacco product regulation*. Paper presented at first Committee Meeting of the Committee on Assessment of Agent-Based Models to Inform Tobacco Product Regulation, Washington, DC.

HHS (U.S. Department of Health and Human Services). 2014. *The health consequences of smoking—50 years of progress: A report of the Surgeon General*. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

Jamal, A., I. T. Agaku, E. O'Connor, B. A. King, J. B. Kenemer, and L. Neff. 2014. Current cigarette smoking among adults—United States, 2005–2013. *Morbidity and Mortality Weekly Report* 63(47):1108–1112.

Lee, B. Y., S. T. Brown, P. C. Cooley, R. K. Zimmerman, W. D. Wheaton, S. M. Zimmer, J. J. Grefenstette, T. M. Assi, T. J. Furphy, D. K. Wagener, and D. S. Burke. 2010. A computer simulation of employee vaccination to mitigate an influenza epidemic. *American Journal of Preventive Medicine* 38(3):247–257.

Manski, C. F. 2013. *Public policy in an uncertain world: Analysis and decisions*. Cambridge, MA: Harvard University Press.

Moore, T. W., P. D. Finley, N. S. Brodsky, T. J. Brown, B. Apelberg, B. Ambrose, R. J. Glass. In press a. Modeling education and advertising with opinion dynamics. *The Journal of Artificial Societies and Social Simulation*.

Moore, T. W., P. D. Finley, B. J. Apelberg, B. Ambrose, N. S. Brodsky, T. J. Brown, C. Husten, R.J. Glass. In press b. An opinion-driven behavioral dynamics model for addictive behaviors. *European Physical Journal B*.

NRC (National Research Council). 2007. *Models in environmental regulatory decision making*. Washington, DC: The National Academies Press.

———. 2012. *Assessing the reliability of complex models: Mathematical and statistical foundations of verification, validation, and uncertainty quantification.* Washington, DC: The National Academies Press.

Orr, M. G., S. Galea, M. Riddle, and G. A. Kaplan. 2014. Reducing racial disparities in obesity: Simulating the effects of improved education and social network influence on diet behavior. *Annals of Epidemiology* 24(8):563–569.

SAMHSA (Substance Abuse and Mental Health Services Administration). 2013. *Results from the 2012 National Survey on Drug Use and Health: Summary of national findings*, NSDUH series H-48, HHS publication No. (SMA) 14-4863. Rockville, MD: U.S. Department of Health and Human Services.

Sanstad, A. H. 2015. Assessing agent-based models for regulatory applications: Lessons from energy analysis. Paper commissioned by the Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation (see Appendix C).

Wagner, W., E. Fisher, and P. Pascual. 2010. Misunderstanding models in environmental and public health regulation. *NYU Law Environmental Law Journal* 18:293–356.

Zhang, D., P. J. Giabbanelli, O. A. Arah, and F. J. Zimmerman. 2014. Impact of different policies on unhealthy dietary behaviors in an urban adult population: An agent-based simulation model. *American Journal of Public Health* 104(7):1217–1222.

1

Introduction

Computational modeling of social processes has been used for many years in numerous disciplines for a variety of purposes, including assisting in the development of public policy decisions. A computational model can be used to inform the regulatory process and can "range from single parameter linear relationship models to models with thousands of separate components and many billions of calculations" (NRC, 2007, p. 36). A growing interest in systems science approaches to population health has led public health researchers, regulators, and others to turn to computational modeling. Many types of models have been used to forecast health effects associated with current and future risk behaviors, including tobacco use. For example, several population dynamics models have been used to simulate the dynamics of smoking use and smoking-attributed deaths in a state or nation and the effects of policies or policy changes on those outcomes (HHS, 2014). Since 2009 the U.S. Food and Drug Administration (FDA) has had broad regulatory authority over tobacco products and has used models as one tool to inform its policy decision-making activities. Recently, FDA has been exploring the usefulness of a particular computational modeling approach—agent-based modeling (ABM)—to inform its policy decisions. (See the section titled Computational Modeling of Tobacco Use on page 25 for more information on ABMs.)

To that end, the Center for Tobacco Products (CTP) at FDA asked the Institute of Medicine (IOM) to review an ABM developed for use by FDA; to comment on its strengths, weaknesses, and usefulness for examining various tobacco regulatory policies; and to provide recommendations on strategies to improve the model and for using ABM in general in the future.

To address that request, the IOM created the Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation (see Box 1-1 and below for a discussion of the committee's statement of task).

At the committee's first meeting, CTP expressed interest in exploring ABMs as a tool for tobacco control policy for several reasons. CTP noted that ABMs have been used to examine complex phenomena and may be particularly useful in providing insight into phenomena for which social interactions and population variation are important factors. CTP explained that ABMs are one tool that might allow CTP to learn more about the importance of individual-level factors that dictate tobacco use, as well as simulate potential use patterns in an evolving market (Fultz, 2014). CTP added that it is motivated by the potential of ABMs to simulate potential effects of policies for which there might be limited data. (See Chapters 3 and 6 for discussion of the limitations of modeling with incomplete data.) CTP also pointed out that ABMs can be helpful when addressing questions where there are ethical issues with using human subjects to conduct the research (Fultz, 2014).

BOX 1-1
Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation
Statement of Task

The Institute of Medicine (IOM) shall convene a committee to assess the applicability of agent-based models of tobacco use and public health as a guide to inform regulators and improve the effect of tobacco regulation policies on public health. The committee shall:

- comment on implications of using agent-based models to examine various tobacco regulatory policies
- assess the strengths and weaknesses of an agent-based model developed for the U.S. Food and Drug Administration (FDA) (to be provided by the Center for Tobacco Products [CTP]) and models currently available in the literature that have been used for similar purposes (to be identified by CTP)
- make recommendations on future directions and strategies to improve the usefulness of the model developed for or to be used by FDA, if needed

ADDRESSING THE STATEMENT OF TASK

The IOM committee was tasked with evaluating a specific ABM commissioned by FDA and to comment more broadly on the application of the ABM approach with respect to tobacco regulatory policy. The committee was also asked to review relevant ABMs in the literature to glean best practices.[1] In addition, CTP asked the committee to identify research gaps related to using ABMs to inform policy (Fultz, 2014).

Because the committee was specifically requested by FDA to evaluate ABMs, that is the major focus of this report. This report addresses modeling techniques similar to ABMs (such as microsimulation), but other potentially useful modeling approaches (such as aggregate models or system dynamics models) are not discussed in detail, except when relevant to ABMs—for example, using ABMs to inform aggregate models. Additionally, it was beyond the scope of this report to discuss when ABMs versus other modeling approaches are suitable to address specific types of questions and contexts. Other reports, however, have compared and contrasted different types of models, including ABMs, and have proposed various ways to identify the appropriateness of using certain modeling approaches for specific situations (Chattoe et al., 2005; Irwin and Wrenn, 2014; NRC, 2014). It is important to note that although some of the discussions in this report are relevant to modeling in general, the assessment of the strengths and limitations of ABMs identified by the committee are not applicable to other modeling approaches unless specified in the report. Furthermore, the committee formally assessed only one ABM in this report as outlined in its statement of task, and although lessons from the development of that model may be applied to future development of ABMs, it is not indicative of the strengths or limitations of other tobacco control ABMs, or tobacco control models using other approaches.

Overview of the Study Process

The IOM convened a 12-member committee (see Appendix E for the committee biographies) with expertise in the fields of modeling, tobacco use behavior and epidemiology, economics, and policy application. To address its charge, the committee gathered information through a variety of means. The committee reviewed literature regarding ABMs, other computational modeling approaches, modeling for policy, and tobacco use behavior. Additionally, the committee heard from various experts in these fields, and

[1] An ABM by Eppstein and colleagues (2011) was specifically identified by CTP for committee review and is discussed in Chapter 4.

explored, learned from, and built on past National Research Council and IOM reports that discuss various modeling techniques, including ABMs.

The committee convened five times between February and November 2014, holding three open-to-the-public information-gathering sessions and two closed-session deliberation meetings. The agendas for the three public meetings can be found in Appendix D. During the first meeting, the committee was presented the charge by CTP as well as the model developed for CTP by Sandia National Laboratories. The second meeting focused on hearing from experts about individual components of the model to be reviewed by the committee, including both its technical and its social and behavioral features. In the third meeting, the committee heard from additional experts and reflected on lessons learned and best practices for using modeling, specifically for informing policy decisions.

The committee received public submissions of materials for its consideration at the meetings and throughout the course of the study.[2] A website was created to provide information to the public about the committee's work and to facilitate communication between the public and the committee.[3] The committee commissioned three experts—Lawrence Blume, Ross Hammond, and Alan Sanstad—to write papers that identify varying views concerning ABM, the practice of and pitfalls associated with ABM, and lessons learned regarding the application of ABMs in health and energy policy. Given the multifaceted approaches to ABMs across disciplines, these papers enriched the committee's discussion and understanding of ABMs from other fields of study and informed the committee's conclusions. As with many fields, there are differences of opinion on how to approach and develop ABMs. These papers provide some of that context and begin to elucidate where there is agreement versus dissention regarding ABMs and identify best practices across varied fields of study. These papers are referenced throughout the report where relevant and are provided in Appendixes A, B, and C.

BACKGROUND

The Continuing Challenge of Tobacco Control

Tobacco consumption continues to be the leading cause of preventable death and disease in the United States (HHS, 2014). More than 42 million Americans, representing approximately 18 percent of the population, currently smoke cigarettes (Agaku et al., 2014; Jamal et al., 2014). Each day, more than 3,200 children under age 18 smoke their first cigarette, and

[2] Public access materials can be requested from: https://www8.nationalacademies.org/cp/ManageRequest.aspx?key=49612.

[3] See http://iom.nationalacademies.org/Reports/2015/Tobacco-Policy-Agent-Based-Modeling.aspx.

more than 700 children become daily cigarette smokers (SAMHSA, 2013). Many of these youth will become addicted and suffer from adverse health consequences. At the current smoking rate, 5.6 million children alive today will die prematurely from smoking-related illness (HHS, 2014).

Tobacco can lead to a wide range of consequences, from debilitating illnesses to severe economic costs. Each year, nearly half a million people in the United States die prematurely from diseases caused by smoking or secondhand smoke exposure, which equates to more than 1,300 deaths every day (HHS, 2014). Life expectancy for smokers is at least 10 years shorter than for nonsmokers (Jha et al., 2013). A 2009 study estimated that U.S. adults have about 14 million major medical conditions that are attributable to smoking (Rostron et al., 2014). Smoking is now associated with 13 types of cancers as well as numerous other diseases, including diabetes and rheumatoid arthritis (HHS, 2014). More than 87 percent of lung cancer deaths, 61 percent of pulmonary disease deaths, and 32 percent of deaths from coronary heart disease are attributable to smoking and exposure to secondhand smoke in the United States (HHS, 2014). In terms of economic burden, tobacco use costs the United States billions of dollars each year. More than $289 billion is incurred in medical expenses and lost productivity from smoking, and $5.6 billion is incurred from lost productivity caused by secondhand smoke (HHS, 2014).

Patterns of tobacco use are evolving (HHS, 2014). Although cigarettes and other combustible products (e.g., cigars, pipes, and hookahs) continue to be the most prevalent forms of adult tobacco use (Agaku et al., 2014), emerging tobacco products, such as electronic cigarettes (e-cigarettes), have been growing in prevalence (Arrazola et al., 2013; HHS, 2014; King et al., 2015). The portion of individuals who had ever used e-cigarettes grew from 3.3 percent to 8.5 percent from 2010 to 2013 for adults over 18 years of age (King et al., 2015), and from 3.3 percent to 6.8 percent from 2011 to 2012 for adolescents in grades 6–12 (Corey et al., 2013). Recent research indicates that the use of e-cigarettes among adolescents has now surpassed the use of traditional tobacco cigarettes or any other tobacco product (Johnston et al., 2015). Furthermore, the use of multiple tobacco products, such as using both smokeless tobacco and cigarettes, has expanded (Apelberg et al., 2014; HHS, 2014; Lee et al., 2014). As new tobacco products rapidly emerge and patterns of tobacco use evolve, the possibility of an increase in initiation and decreased or delayed cessation among youth and young adults is cause for concern. Simulation models have been used by FDA to help address the critical health and social concerns of the present smoking epidemic.

Overview of FDA's Authority Over Tobacco Products

Until 2009, tobacco products were exempt from regulation under the nation's federal health and safety laws. FDA had regulated food, drugs (including nicotine replacements), and cosmetics for many decades, but not tobacco products. On June 22, 2009, the Family Smoking Prevention and Tobacco Control Act (Tobacco Control Act) gave FDA the authority to regulate the manufacture, distribution, and marketing of tobacco products to protect public health and reduce tobacco use in the United States. To oversee the implementation of the law, FDA established CTP. The goals of CTP are to prevent people from starting to use tobacco products, encourage current tobacco users to quit, and reduce the harm caused by tobacco use.

The Tobacco Control Act gives FDA the authority to regulate cigarettes, cigarette tobacco, roll-your-own tobacco, and smokeless tobacco. Additional tobacco products, such as e-cigarettes and cigars, are being considered through a deeming proposal.[4] The Tobacco Control Act gives FDA the authority, through rule making, to adopt tobacco product standards appropriate for the protection of public health. FDA can adopt new product standard provisions to reduce addiction, reduce toxicity and carcinogenicity, reduce harmful constituents, restrict sale and distribution, and address the form and content of labeling for the proper use of tobacco products. Other authorities include restricting advertising and promotion and imposing the placement of health warnings on products. (See Chapter 2 for a detailed discussion of FDA's regulatory authority.)

New tobacco policies and regulations must be based on available medical, scientific, and other technological evidence as appropriate for the protection of the public health. In particular, FDA reviews new tobacco products on the basis of a public health standard instead of the "safe and effective" standard that it uses to evaluate drugs. The public health standard requires FDA to consider scientific evidence concerning (1) the risks and benefits to the general public, including users and nonusers of tobacco products; (2) the increased or decreased likelihood that existing users of tobacco products will stop using the products; and (3) the increased or decreased likelihood that those who do not use tobacco products will start using the new products. FDA considers the net effect on tobacco-related behavior changes within the whole population for initiation, cessation, and relapse. Consequently, FDA is concerned with effectively forecasting the public health effects of potential changes in tobacco standards and other policies.

[4]The Tobacco Control Act gives FDA the ability to regulate other tobacco products through rule making. In early 2014, FDA proposed a "deeming" rule that would extend the agency's authority to cover other products that meet the definition of a tobacco product, such as e-cigarettes, cigars, pipe tobacco, waterpipe (hookah) tobacco, and nicotine gels and dissolvables. If passed, FDA would be able to regulate these newly deemed products in ways consistent with currently regulated tobacco products.

COMPUTATIONAL MODELING OF TOBACCO USE

Computational modeling is among the many tools that can be used to inform and evaluate tobacco control policies. In the past, population-based aggregate models and microsimulations of tobacco control have been used to model the effect of tobacco control policies, trends in smoking prevalence, and health outcomes (HHS, 2014; Holford et al., 2014; Levy et al., 2006; Mendez et al., 1998; Moolgavkar et al., 2012; Orme et al., 2001; Tengs et al., 2001). (See Box 1-2 for a brief summary of tobacco control modeling efforts to date; for a more detailed overview see Appendix 15.1 of the 2014 Surgeon General's report.) Currently, FDA is exploring other modeling approaches, including ABMs, to forecast effectively the public health effects of potential changes in tobacco standards and other policies. ABM is a type of computer simulation that studies complex systems by exploring how individual elements (agents) of a system behave as a function of individual characteristics, and interactions with each other, and with the environment. Each agent interacts with other agents based on a set of rules and within an environment specified by the modeler; these interactions lead to a set of specific outcomes, some of which may be unexpected. (See Chapter 3 and Appendix A for a detailed discussion of ABMs.) Because ABMs can be used to explore the potential impact of policies and interventions in dynamic social and physical environments, they may be a useful tool to aid in decision making among policy makers. ABMs have been used to examine other public health interventions and policies, such as for infectious diseases (Epstein et al., 2007; Lee et al., 2010) and obesity (Auchincloss et al., 2011; Orr et al., 2014; Zhang et al., 2014), but the use of ABMs has not been fully explored and considered in the tobacco regulatory space (Hammond, 2015).

It should not be inferred that the committee or FDA found that existing models are not useful. Researchers and policy makers have used existing tobacco control models extensively to inform policy decisions and those models continue to be a useful and important tool. This report is meant to grow on the large body of work on tobacco control modeling by exploring how ABMs might be a helpful tool to add to the existing modeling toolkit (see the section titled Why Use Agent-Based Models to Explore Tobacco Use? on page 52 for a discussion on the role of ABMs for tobacco regulation).

Center for Tobacco Products and Agent-Based Modeling

Through an interagency agreement between CTP and the U.S. Department of Energy, CTP commissioned Sandia National Laboratories[5] to

[5] The modeling team is part of the Complex Adaptive System of Systems (CASoS) Engineering Initiative.

BOX 1-2
Brief Overview of Tobacco Control Models

Over the past few decades, many types of models have been used to inform tobacco control research and policy (Chaloupka and Warner, 2000; HHS, 2014). For example, aggregate (also called compartmental or population) models have been used for over 15 years in tobacco control (Holford et al., 2014; Levy et al., 2006; Warner and Mendez, 2012). These models simulate the evolution of populations between non-overlapping categories (e.g., nonsmokers, current smokers, and former smokers). The evolution is dictated by rates either built in or used as inputs for the model, such as initiation and cessation rates and birth and death rates. The models can be used to predict outcomes assuming no change in current smoking rates and trends or else assuming estimated changes in smoking rates and trends. They are also useful for comparing rates and trends after a policy has been implemented with a set of counterfactual data (that is, a comparison between what actually happened and what would have happened in the absence of the intervention) (Holford et al., 2014).

Some tobacco control aggregate models have been used to examine the effects of policies on smoking prevalence, cessation, quit attempts, and other measures of tobacco use in a population. Several of these model scenarios were initially validated with national-level data and have since been adapted to look at individual states or other countries. Examples of tobacco control aggregate models include the University of Michigan Tobacco Prevalence and Health Effects Model and the SimSmoke model, which have been used to look at potential policies, such as an analysis carried out for FDA of the ramifications of a menthol ban. Other models have explored the cost-effectiveness of smoking interventions (BENESCO model), the health-related economic impact of smoking (SAMMEC model), and various health outcomes from smoking (CANSAVE, CISNET's [Cancer Intervention and Surveillance Modeling Network's] Yale Lung Cancer model) (HHS, 2014).

Microsimulations, which model at the individual level, have also been used to study tobacco control. These models have quantified the impact of tobacco control on lung cancer mortality and smoking- related mortality in the United States over the past few decades, contributing to the advancement of lung cancer screening strategies and public health research (de Koning et al., 2014; McMahon et al., 2014; Moolgavkar et al., 2012). For example, to model the natural history of lung cancer, six independent microsimulation models were developed as part of CISNET (McMahon et al., 2012).

develop an ABM that could help FDA understand the potential impacts of a variety of policies on population health. The modeling efforts by Sandia National Laboratories under this agreement include population health models that aim to help forecast potential long-term impacts on prevalence, morbidity, and mortality for the population in the United States and various other types of models, including an ABM.

The ABM being developed by Sandia National Laboratories for FDA is called Social Network Analysis for Policy on Directed Graph Networks, or SnapDragon (Moore et al., in press a,b). The main purpose of the model is to explore the effects of various tobacco policies and interventions such as public education campaigns on opinion and tobacco use within social networks. At the time of this review, SnapDragon was still in an early development stage. The authors have presented the model at professional meetings, but as of publication of this report, no peer-reviewed papers on Snapdragon have been published. (See Chapter 5 for a detailed review of the SnapDragon model.)

Modeling and Policy

Models are used to inform regulatory policy for several reasons. They can better describe complex and poorly characterized problems, but are not "truth generating machines" (NRC, 2007, p. 182). Although policy-relevant computational models are incomplete representations of a small piece of the regulatory environment, this does not mean they lack value. They can provide "other assets to policymaking, such as providing a conceptual map of existing relationships, highlighting new interconnections, and elucidating important uncertainties, all of which significantly aid policy deliberation, but do not replace it" (Wagner et al., 2010, p. 295). These models can also build theory inductively or deductively, or both; guide future data collection by pinpointing unknowns and seeing which appear to matter; explicitly inform intervention design by anticipating consequences; and integrate data that are scattered across different sources and use the interaction of the data informatively (Epstein, 2008). Because of these capabilities, models can be used as one piece of the evidence base to inform the design of future policies, evaluate the effects of current or past policies, and identify key leverage points and opportunities for policy making.[6] As will be discussed later in this report, to inform policy effectively, policy makers need to understand the level of model uncertainty, what the model does and does not forecast, to what extent the model is suitable for the question or process under study, and how the model fits within the body of available evidence.

Report Contents

To address its statement of task, the committee reviews and discusses the complex environment in which tobacco control policies are created and how ABMs could be a useful tool to assist in tobacco control policy deci-

[6] See Appendix A for a comprehensive discussion on using ABMs to inform policy.

sions (Chapter 2), reviews the structure and implications of using ABMs to inform policy decisions (Chapter 3), develops evaluation criteria for the review of ABMs (Chapter 4), and illustrates the evaluation criteria with an evaluation of the SnapDragon model (Chapter 5). Where applicable, parts of this evaluation framework are exemplified by reviewing other relevant models to illustrate how the framework is used. This report focuses on tobacco policies that fall within the realm of FDA's purview, but it is not intended to discourage or ignore modeling or related policy efforts by others, such as other governmental and nongovernmental organizations, states, localities, or other policy scientists, but rather to best address the current needs of the report sponsor. In Chapter 6 the committee discusses inputs for ABMs for tobacco control, and makes recommendations for future implementation of ABMs at CTP.

REFERENCES

Agaku, I. T., B. A. King, C. G. Husten, R. Bunnell, B. K. Ambrose, S. S. Hu, E. Holder-Hayes, H. R. Day, and Centers for Disease Control and Prevention. 2014. Tobacco product use among adults—United States, 2012–2013. Morbidity and Mortality Weekly Report 63(25):542–547.

Apelberg, B. J., C. L. Backinger, and S. J. Curry. 2014. Enhancing youth tobacco surveillance to inform tobacco product regulation: Findings from the 2012 National Youth Tobacco Survey. *American Journal of Preventive Medicine* 47(2 Suppl 1):S1–S3.

Arrazola, R. A., S. Dube, and B. A. King. 2013. Tobacco product use among middle and high school students—United States, 2011 and 2012. *Morbidity and Mortality Weekly Report* 62(45):893–897.

Auchincloss, A. H., R. L. Riolo, D. G. Brown, J. Cook, and A. V. Diez Roux. 2011. An agent-based model of income inequalities in diet in the context of residential segregation. *American Journal of Preventive Medicine* 40(3):303–311.

Chaloupka, F. J., and K. E. Warner. 2000. The economics of smoking. In *Handbook of health economics edition 1*. Vol. 1, edited by A. J. Culyer and J. P. Newhouse. Amsterdam, Netherlands: North Holland, Elsevier. Pp. 1539–1627.

Chattoe, E., M. Hickman, and P. Vickerman. 2005. *Drugs futures 2025? Modelling drug use*. London, UK: DTI Foresight, Office of Science and Technology. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/299223/05-1183-drugs-futures-modelling.pdf (accessed March 30, 2015).

Corey, C. G., B. Wang, S. E. Johnson, B. J. Apelberg, C. Husten, B. A. King, T. McAfee, R. Bunnell, R. Arrazola, S. R. Dube, and Centers for Disease Control and Prevention. 2013. Electronic cigarette use among middle and high school students—United States 2011–2012. *Morbidity and Mortality Weekly Report* 62(35):729–730.

de Koning, H. J., R. Meza, S. K. Plevritis, K. ten Haaf, V. N. Munshi, J. Jeon, S. A. Erdogan, C. Y. Kong, S. S. Han, J. van Rosmalen, S. E. Choi, P. F. Pinsky, A. B. de Gonzalez, C. D. Berg, W. C. Black, M. C. Tammemägi, W. D. Hazelton, E. J. Feuer, and P. M. McMahon. 2014. Benefits and harms of computed tomography lung cancer screening strategies: A comparative modeling study for the U.S. Preventive Services Task Force. *Annals of Internal Medicine* 160(5):311–320.

Eppstein, M. J., D. K. Grover, J. S. Marshall, and D. M. Rizzo. 2011. An agent-based model to study market penetration of plug-in hybrid electric vehicles. *Energy Policy* 39(6):3789–3802.

Epstein, J. M. 2008. Why model? *Journal of Artificial Societies & Social Simulation* 11(4):12.

Epstein, J. M., D. M. Goedecke, F. Yu, R. J. Morris, D. K. Wagener, and G. V. Bobashev. 2007. Controlling pandemic flu: The value of international air travel restrictions. *PLoS ONE* 2(5):e401.

Fultz, S. L. 2014. *Committee on assessment of agent-based models to inform tobacco product regulation*. Paper presented at First Committee Meeting of the Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation, February 26, 2014, Washington, DC.

Hammond, R. A. 2015. Considerations and best practices in agent-based modeling to inform policy. Paper commissioned by the Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation (see Appendix A).

HHS (U.S. Department of Health and Human Services). 2014. *The health consequences of smoking—50 years of progress: A report of the Surgeon General*. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

Holford, T. R., R. Meza, K. E. Warner, C. Meernik, J. Jeon, S. Moolgavkar, and D. T. Levy. 2014. Tobacco control and the reduction in smoking-related premature deaths in the United States, 1964–2012. *JAMA* 311(2):164–171.

Irwin, E. G., and D. H. Wrenn. 2014. An assessment of empirical methods for modeling land use. In *The Oxford handbook of land economics*, edited by J. M. Duke and J. Wu. New York: Oxford University Press.

Jamal, A., I. T. Agaku, E. O'Connor, B. A. King, J. B. Kenemer, and L. Neff. 2014. Current cigarette smoking among adults—United States, 2005–2013. *Morbidity and Mortality Weekly Report* 63(47):1108–1112.

Jha, P., C. Ramasundarahettige, V. Landsman, B. Rostron, M. Thun, R. N. Anderson, T. McAfee, and R. Peto. 2013. 21st-century hazards of smoking and benefits of cessation in the United States. *New England Journal of Medicine* 368(4):341–350.

Johnston, L. D., P. M. O'Malley, R. A. Miech, J. G. Bachman, and J. E. Schulenberg. 2015. *Monitoring the future national results on adolescent drug use: 1975–2014: Overview of key findings on adolescent drug use*. Ann Arbor, MI: Institute for Social Research, University of Michigan.

King, B. A., R. Patel, K. Nguyen, and S. R. Dube. 2015. Trends in awareness and use of electronic cigarettes among U.S. Adults, 2010–2013. *Nicotine & Tobacco Research* 17(2):219–227.

Lee, B. Y., S. T. Brown, P. C. Cooley, R. K. Zimmerman, W. D. Wheaton, S. M. Zimmer, J. J. Grefenstette, T. M. Assi, T. J. Furphy, D. K. Wagener, and D. S. Burke. 2010. A computer simulation of employee vaccination to mitigate an influenza epidemic. *American Journal of Preventive Medicine* 38(3):247–257.

Lee, Y. O., C. J. Hebert, J. M. Nonnemaker, and A. E. Kim. 2014. Multiple tobacco product use among adults in the United States: Cigarettes, cigars, electronic cigarettes, hookah, smokeless tobacco, and snus. *Preventive Medicine* 62:14–19.

Levy, D. T., J. E. Bauer, and H. R. Lee. 2006. Simulation modeling and tobacco control: Creating more robust public health policies. *American Journal of Public Health* 96(3):494–498.

McMahon, P. M., W. D. Hazelton, M. Kimmel, and L. D. Clarke. 2012. Chapter 13: CISNET lung models: Comparison of model assumptions and model structures. *Risk Analysis* 32(Suppl 1):S166–S178.

McMahon, P. M., R. Meza, S. K. Plevritis, W. C. Black, C. M. Tammemägi, A. Erdogan, K. ten Haaf, W. Hazelton, T. R. Holford, J. Jeon, L. Clarke, C. Y. Kong, S. E. Choi, V. N. Munshi, S. S. Han, J. van Rosmalen, P. F. Pinsky, S. Moolgavkar, H. J. de Koning, and E. J. Feuer. 2014. Comparing benefits from many possible computed tomography lung cancer screening programs: Extrapolating from the national lung screening trial using comparative modeling. *PLoS ONE* 9(6):e99978.

Mendez, D., K. E. Warner, and P. N. Courant. 1998. Has smoking cessation ceased?: Expected trends in the prevalence of smoking in the United States. *American Journal of Epidemiology* 148(3):249–258.

Moolgavkar, S. H., T. R. Holford, D. T. Levy, C. Y. Kong, M. Foy, L. Clarke, J. Jeon, W. D. Hazelton, R. Meza, F. Schultz, W. McCarthy, R. Boer, O. Gorlova, G. S. Gazelle, M. Kimmel, P. M. McMahon, H. J. de Koning, and E. J. Feuer. 2012. Impact of reduced tobacco smoking on lung cancer mortality in the United States during 1975–2000. *Journal of the National Cancer Institute* 104(7):541–548.

Moore, T. W., P. D. Finley, N. S. Brodsky, T. J. Brown, B. Apelberg, B. Ambrose, R. J. Glass. In press a. Modeling education and advertising with opinion dynamics. *The Journal of Artificial Societies and Social Simulation.*

Moore, T. W., P. D. Finley, B. J. Apelberg, B. Ambrose, N. S. Brodsky, T. J. Brown, C. Husten, R.J. Glass. In press b. An opinion-driven behavioral dynamics model for addictive behaviors. *European Physical Journal B*.

NRC (National Research Council). 2007. *Models in environmental regulatory decision making*. Washington, DC: The National Academies Press.

———. 2014. *Advancing land change modeling: Opportunities and research requirements*. Washington, DC: The National Academies Press.

Orme, M. E., S. L. Hogue, L. M. Kennedy, A. C. Paine, and C. Godfrey. 2001. Development of the health and economic consequences of smoking interactive model. *Tobacco Control* 10(1):55–61.

Orr, M. G., S. Galea, M. Riddle, and G. A. Kaplan. 2014. Reducing racial disparities in obesity: Simulating the effects of improved education and social network influence on diet behavior. *Annals of Epidemiology* 24(8):563–569.

Rostron, B. L., C. M. Chang, and T. F. Pechacek. 2014. Estimation of cigarette smoking–attributable morbidity in the United States. *JAMA Internal Medicine* 174(12):1922–1928.

SAMHSA (Substance Abuse and Mental Health Services Administration). 2013. *Results from the 2012 National Survey on Drug Use and Health: Summary of national findings*, NSDUH series H-48, HHS publication No. (SMA) 14-4863. Rockville, MD: U.S. Department of Health and Human Services.

Tengs, T. O., N. D. Osgood, and T. H. Lin. 2001. Public health impact of changes in smoking behavior: Results from the tobacco policy model. *Medical Care* 39(10):1131–1141.

Wagner, W., E. Fisher, and P. Pascual. 2010. Misunderstanding models in environmental and public health regulation. *NYU Law Environmental Law Journal* 18:293–356.

Warner, K. E., and D. Mendez. 2012. Accuracy and importance of projections from a dynamic simulation model of smoking prevalence in the United States. *American Journal of Public Health* 102(11):2045–2048.

Zhang, D., P. J. Giabbanelli, O. A. Arah, and F. J. Zimmerman. 2014. Impact of different policies on unhealthy dietary behaviors in an urban adult population: An agent-based simulation model. *American Journal of Public Health* 104(7):1217–1222.

2

Tobacco Control Landscape

Given that tobacco smoking is the leading preventable cause of death in the United States, reducing smoking rates is one of the nation's highest public health priorities. Doing so would lower deaths and illness caused by tobacco products, decrease related health care costs, and improve the quality and length of life of individuals. Although great strides have been made since the landmark Surgeon General's report in 1964, almost 18 percent of the U.S. population still smokes (Agaku et al., 2014; Jamal et al., 2014). The environment in which tobacco products are used and sold is complex and evolving, and requires working across multiple sectors and understanding an intricate web of stakeholders including a diverse user population, largely addicted to tobacco products. Understanding this tobacco landscape is essential when attempting to model tobacco control policies to inform policy decision making. This chapter provides a snapshot of these issues, including policy inputs and context that will be reviewed in this report, a review of what is currently understood about tobacco use behavior, and why agent-based models (ABMs) could be a useful tool for exploring tobacco-related questions.

TOBACCO ENVIRONMENT

The tobacco regulatory environment is complicated, and when the intricate web of other actors (or agents) is considered, tobacco regulation can be viewed as a complex adaptive system. Modeling is a useful tool for understanding the structure and behavior of complex adaptive systems, which are defined by Plsek and Greenhalgh (2001, p. 625) as "a collection of individual agents with freedom to act in ways that are not always totally predictable, and whose actions are interconnected so that one agent's actions changes the context for other agents." Components of these systems can be studied separately, but there are relationships among them, so the behavior of each component depends on the behavior of others. Below the committee highlights some of the complex relationships among various stakeholders that are particularly relevant to the U.S. Food and Drug Administration (FDA) today. See Figures 2-1 and 2-2 for depictions of these relationships, highlighting different aspects of the complex tobacco system. Figure 2-1 illustrates the web of relationships in the tobacco environment, which consists of varied agents in the system, including regulators, tobacco retailers, the tobacco control community, health care, and others. Figure 2-2 displays the many feedbacks that operate among various groups and entities within the tobacco control environment, including tobacco research, individual behavior, tobacco control programs, tobacco industry, and economics.

The Complex Tobacco Problem

Tobacco has been referred to as a wicked problem (APSC, 2007; Dorfman and Wallack, 1993; Young et al., 2012). Originally coined by Rittel and Webber (1973), a wicked problem has more recently been defined as "a social or cultural problem that is difficult or impossible to solve for as many as four reasons: incomplete or contradictory knowledge, a large number of people and opinions involved, the important economic burden, and the interconnected nature of these problems with other problems" (Kolko, 2012). Wicked problems can productively be viewed within the environment of the larger complex adaptive system, which describes the wider landscape that surrounds and influences the problem. Wicked problems cannot be approached solely with analytical approaches, nor can they be managed by a single organization, jurisdiction, or domain (Young et al., 2012). Classic examples of wicked problems include poverty, climate change, and land degradation; the problem of obesity is an example of such a problem that has arisen more recently. Tobacco is viewed as a wicked problem because of the often contradictory goals of stakeholders that give rise to uncertainty and because of the addictive nature of tobacco products.

Five years ago, FDA was given an unprecedented opportunity to regulate tobacco, but the complex nature of tobacco control remains an impediment to clear-cut and effective policy implementation. (See later in this chapter for a discussion of FDA's regulatory authorities.) FDA's steps toward comprehensive tobacco regulation have been gradual for a number of reasons. The Family Smoking Prevention and Tobacco Control Act (Tobacco Control Act) was enacted only 5 years ago, so the agency is still developing its strategy and focus. One of the first—and key—regulatory

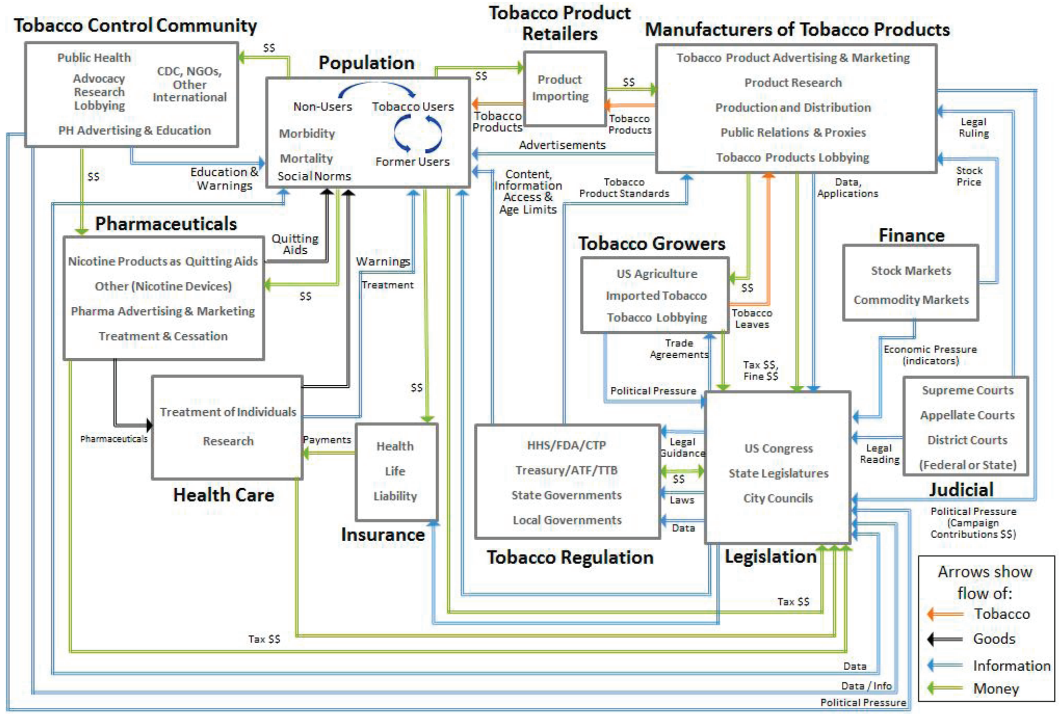

FIGURE 2-1 Agents and relationships within the tobacco control landscape.
NOTES: ATF = Bureau of Alcohol, Tobacco, Firearms and Explosives; CDC = Centers for Disease Control and Prevention; CTP = Center for Tobacco Products; FDA = U.S. Food and Drug Administration; HHS = U.S. Department of Health and Human Services; NGO = nongovernmental organization; PH = public health; TTB = Alcohol and Tobacco Tax and Trade Bureau.
SOURCE: SNL, 2014.

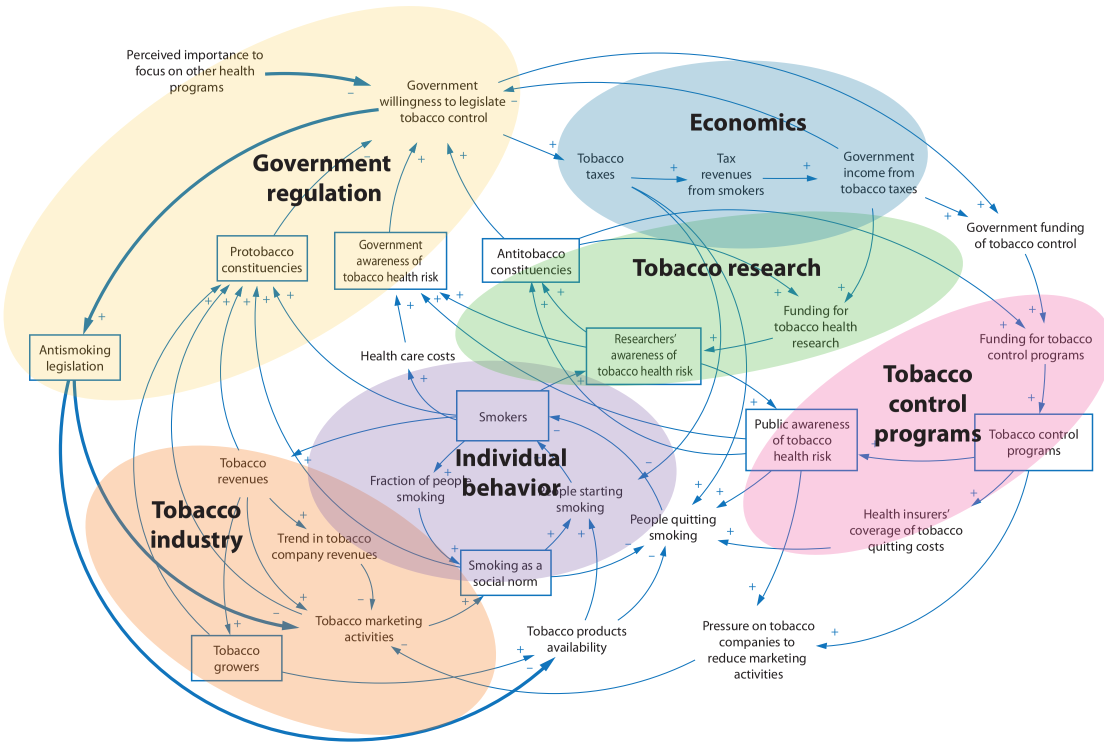

FIGURE 2-2 Complex tobacco landscape.
NOTE: This figure is not drawn to scale, nor is any meaning implied by the relative sizes of elements within the figure. See Kirkwood (1998) for more information about causal loop diagram construction.
SOURCE: NCI, 2007, adapted by Luke, 2013.

efforts, the move to put graphic warning labels on packs of cigarettes, was at least temporarily thwarted by the tobacco industry.[1] Other facets of FDA's regulatory authority over tobacco, such as reducing the nicotine content of regulated products and banning menthol, remain largely untested. FDA is also limited as to how and what it can regulate based on the act. For instance, it cannot eliminate nicotine content in tobacco products or regulate products geared toward cessation, and some decisions, such as tobacco retailer density, are made at the state or local level.

Furthermore, as is the case with other federal agencies, FDA has congressional oversight accompanied by ongoing discussions regarding modification of the Tobacco Control Act (NACS, 2014). In addition, the regulatory process requires considerable public input, which means that any proposed regulation must be posted for public comment for a period that is generally between 30 and 60 days but can, in some cases, be longer (OFR, 2014). Because many policies promulgated by the agency are likely to be brought to court, FDA must often plan for at least the risk of legal challenges (TCLC, 2014c). In addition to the case of graphic warning labels, FDA has already faced major court challenges. In 2010, in the case *Sottera, Inc. v. FDA*,[2] a federal court ruled against FDA's attempt to expand its authority to electronic cigarettes by deeming them a medical drug-device, which would fall under FDA's jurisdiction under the Federal Food, Drug, and Cosmetic Act (FD&C Act).[3] More recently, a suit was filed by Lorillard and R.J. Reynolds challenging a CTP menthol report (TPSAC, 2011) on the grounds that several members of the FDA's Tobacco Products Scientific Advisory Committee (TPSAC), many of whom contributed to the report, had conflicts of interest. In July 2014, a District Court judge ruled that CTP could not use the 2011 report and that the membership of TPSAC should be reconstituted.[4] FDA has filed a notice of appeal to the U.S. Court of Appeals for the District of Columbia Circuit (TCLC, 2014c).

As tobacco products continue to be introduced, FDA and other members of the tobacco control community face evolving challenges and uncertainty. The tobacco control community includes other federal agencies, such

[1] Five tobacco manufacturers filed suit in the U.S. District Court for the District of Columbia to challenge the FDA's final regulation governing graphic warning labels for cigarettes on August 16, 2011. The court found that the graphic warning rule unconstitutionally limited the tobacco companies' right to freedom of speech (*R.J. Reynolds Tobacco Co. v. United States Food & Drug Admin.*, 845 F.Supp.2d 266 (D.D.C. 2012)). On appeal, the U.S. Court of Appeals for the DC Circuit upheld the district court's finding that the graphic warning requirement was unconstitutional (*R.J. Reynolds Tobacco Co. v. United States Food & Drug Admin.*, 696 F.3d 1205 (D.C. Cir. 2012)).

[2] *Sottera Inc v. FDA*, 627 F.3d 891 (D.C. Cir. 2010).

[3] Federal Food, Drug, and Cosmetic Act ("FD&C Act"), 21 U.S.C. § 351 et seq.

[4] *Lorillard, Inc., et al., v. FDA, et al.,* No. 11-440 (D.C. Cir. 2014).

as the Centers for Disease Control and Prevention (CDC), and advocacy and nonprofit organizations such as Legacy and Campaign Tobacco Free Kids. Although the tobacco control community has a general mission to curb the tobacco epidemic, actors within the tobacco control community have different mechanisms with which to combat tobacco use. Advocacy groups, for example, can use political pressure to fight to reduce tobacco use. These various types of groups sometimes disagree regarding the most effective approach to combat tobacco use, particularly when there are limited data and uncertainty in research findings. Disagreement among experts can delay policy making. Electronic cigarettes (e-cigarettes) are a relevant example. Some public health officials argue that e-cigarettes may help smokers quit smoking or reduce harm by encouraging smokers to smoke fewer cigarettes (Abrams, 2014; Fairchild et al., 2014; Hajek, 2013). But others are concerned that the use of e-cigarettes may result in delayed or deferred quitting, renormalize smoking behaviors, and lead to the continuing sales of conventional cigarettes (Bhatnagar et al., 2014; Grana et al., 2014).

Tobacco growers, manufacturers of tobacco products, and tobacco product retailers continue to facilitate the production, distribution, and promotion of tobacco products. To counter antitobacco efforts, tobacco-manufacturing groups have invested billions in advertisements and promotions, including payments to retailers and pharmacies.[5] Tobacco growers and manufacturing groups have leveraged political ties through campaign contributions and lobbying and brought lawsuits to block antitobacco legislation (Morley et al., 2002). They have also collaborated with other industries, such as the hospitality industry, to challenge comprehensive clean indoor air laws (Traynor et al., 1993), and with financial analysts from the investment bank industry to promote the tobacco industry's public policy agenda (Alamar and Glantz, 2004; Tsoukalas and Glantz, 2003). As the market for alternative tobacco products grows, new stakeholders, such as e-cigarette companies and trade associations, have emerged. The Smoke Free Alternatives Trade Association, for example, engages lobbyists at the federal and state levels to block potentially threatening legislation related to vapor products and aims to reinforce the distinctions between vapor and tobacco products and their two respective industries (SFATA, 2015). Meanwhile, tobacco manufacturers have acquired e-cigarette companies and have test marketed their own e-cigarette products (Bauld et al., 2014).

Many other players and infrastructure are linked to—and, in some cases, dependent on—tobacco products. Retail, drug, and vape stores as well as shipping and marketing companies can derive considerable profit from tobacco products. Furthermore, state and local governments often

[5] An exception: in September 2014, CVS/Caremark was the first pharmacy to stop selling tobacco products.

benefit considerably from tobacco taxes and also receive tobacco settlement funds in varying degrees (NAAG, 1998), thus mitigating their incentive to eliminate tobacco use. Pharmaceutical companies that develop and sell cessation products may also be concerned with a cut in profits if dramatic tobacco reductions occur.

Finally, there are the 42 million current smokers in the United States (HHS, 2014b). Individuals start, maintain, and stop the use of tobacco for a variety of reasons. For example, adolescents may experiment with smoking because of peer pressure or social norms, whereas others may use tobacco regularly to relieve stress. Some demographic groups have particularly high smoking prevalence. Adults with mental illness, for instance, are approximately twice as likely to smoke as those who have not been diagnosed with mental illness (Gfroerer et al., 2013; Lasser et al., 2000), potentially to self-medicate, regulate moods, and mitigate stress (Ziedonis et al., 2008). Smoking prevalence among lesbian, gay, bisexual, and transgender individuals in the United States is also significantly higher than in the general population (Fallin et al., 2014; King et al., 2012; Lee et al., 2009); this may be due to stigma, discrimination, and stress (Blosnich et al., 2013; Hatzenbuehler et al., 2011) or to being targeted by tobacco industry marketing efforts (Dilley et al., 2008; Stevens et al., 2004; Washington, 2002) or by media (Lee et al., 2013; Smith et al., 2006), or other reasons (Pokhrel and Caine, 2012). Whatever the reason for using tobacco products, many of these smokers are likely to become addicted, specifically to nicotine. However, addiction affects individuals differently depending on several factors, such as the age of initiation and genetic susceptibilities, meaning individuals may differ in their abilities to quit (Benowitz, 2008a). Furthermore, barriers to cessation treatments and support services that may be experienced by some populations, such as ethnic minorities and people from low socioeconomic backgrounds, may reduce the success rates of quitting (Fu et al., 2007; Malarcher et al., 2011; TUDGP, 2008). The continual introduction of new tobacco products in the market, such as electronic nicotine delivery devices, adds another layer of complexity.

This brief summary provides only a glimpse of the complex tobacco landscape. The multiple players, interdependencies, unintended side effects, and contradictions in an evolving environment that make tobacco a wicked problem complicate the process of finding and implementing ways to reduce tobacco use. Even when a solution seems to have been found, unintended consequences may emerge. In effect, creative methods that can anticipate alternative scenarios and unforeseen consequences could be useful. These considerations lead to the use of analytical methods to understand the current policy options and the effects of those policies and to predict the effect of potential new tobacco regulatory policies. ABMs may be one such method. To understand how major stakeholders in tobacco control, such as

FDA, could use and gain from modeling, it is important to further examine the regulatory approaches and potential policies for tobacco products. (See the final section of this chapter for a discussion on how ABMs could be useful tools for informing tobacco control policy.)

REGULATORY AUTHORITY FOR TOBACCO PRODUCTS

Across the United States there are many policies and interventions in place meant to reduce tobacco use either by preventing initiation or by encouraging cessation. These policies have been put in place by federal, state, and local governments. This report focuses on the policies or interventions under the purview of FDA. However, actions by others in the tobacco regulatory environment can affect policies put forth by FDA, so these are briefly outlined in this chapter as well.

Federal Regulation of Tobacco Products

U.S. Food and Drug Administration

As of 2009, the Family Smoking Prevention and Tobacco Control Act[6,7] (the Tobacco Control Act) gave FDA (part of the U.S. Department of Health and Human Services [HHS]) broad authority to regulate the manufacturing, marketing, and sale of tobacco products, including cigarettes, cigarette tobacco, "roll-your-own" tobacco, and smokeless tobacco products (see Box 2-1 for highlights from the Tobacco Control Act).[8] Recently, FDA proposed regulations to extend its authority to regulate other tobacco products. FDA oversees the implementation of the Tobacco Control Act through a variety of mechanisms. For example, FDA developed CTP to implement TPSAC to provide advice, information, and recommendations to FDA. Additionally, FDA assesses and collects user fees from tobacco product manufacturers and importers based on their market share, and uses the money to fund FDA activities related to the regulation of tobacco

[6] Family Smoking Prevention and Tobacco Control Act of 2009, Public Law 111-31, 111th Cong. (June 22, 2009).

[7] The Food, Drug, and Cosmetic Act was expanded to include the Tobacco Control Act; see Subchapter IX—Tobacco Products (sections 387–387u).

[8] Two other major tobacco acts preceded the Tobacco Control Act. The Federal Cigarette Labeling and Advertising Act (FCLAA) of 1966, 15 U.S.C. § 1335a(a), Public Law 89-92, was amended by the Comprehensive Smoking Education Act (CSTHEA) of 1986, 15 U.S.C. § 4401-4408, Public Law 99-252 (February 27, 1986). CSTHEA, as amended by the 2009 Tobacco Control Act, "requires manufacturers, packagers, and importers of cigarettes to place one of four statutorily-prescribed health-related warnings on cigarette packages and in advertisements, on a rotating basis." CSTHEA prohibits any advertising of smokeless tobacco products on radio and television.

BOX 2-1
Highlights from the
Family Smoking Prevention and Tobacco Control Act

- Seeks to prevent and reduce tobacco use by young people
- Recognizes that tobacco products are legal products available for adult use
- Prohibits false or misleading labeling and advertising for tobacco products
- Allows FDA to establish product standards and to require scientific evidence for any claims of reduced exposure and harm
- Provides the tobacco industry with some mechanisms to submit an application to FDA for new products or tobacco products with modified risk claims
- Grants FDA enforcement authority and general set of sanctions for violations of the law and allows FDA to contract with states to support FDA with retailer inspections

SOURCE: FDA, 2014.

products.[9] FDA also issues regulations and conducts inspections to investigate illicit trade in tobacco products.[10] A detailed list of what the FDA does and does not have authority over is described below in five categories: manufacturing, marketing, sales, limitations, and new tobacco products. (The following sections on manufacturing, marketing, and sales are largely excerpted from FDA, 2014.)

Manufacturing FDA has the authority to oversee several areas regarding manufacturing, including (FDA, 2014):

- Registration and inspection of tobacco companies
  - Requiring owners and operators of tobacco companies to register annually and be subject to inspection every 2 years by FDA
- Standards for tobacco products
  - Allowing FDA to require standards for tobacco products (e.g., tar and nicotine levels) as appropriate to protect public health

[9]The Tobacco Control Act user fee program will generate more than $4.5 billion in user fees over 9 years (2009–2018) (FDA, 2009). FDA spent (obligated) less than half of the $1.1 billion in tobacco user fees it collected from manufacturers and others from fiscal year 2009 through the end of fiscal year 2012 (Crosse, 2014).

[10]For more information on the illicit tobacco market in the United States, see NRC (2015), *Understanding the U.S. Illicit Tobacco Market: Characteristics, Policy Context, and Lessons from International Experiences*.

  - Banning cigarettes with characterizing flavors, except menthol and tobacco
- "Premarket Review" of new tobacco products
  - Requiring manufacturers who wish to market a new tobacco product to obtain a marketing order from FDA prior to marketing that new product[11]
- Modified risk products
  - Requiring manufacturers who wish to market a tobacco product with a claim of reduced harm to obtain a marketing order from FDA
- Requiring tobacco companies to disclose research on the health, toxicological, behavioral, or physiologic effects of tobacco use
  - Requiring tobacco companies to disclose information on ingredients and constituents in tobacco products and to notify FDA of any changes

Marketing FDA has some authority related to tobacco product advertising aimed at youth, the use of certain claims regarding tobacco, use of warning labels, and enforcement of policies made in this area, including (FDA, 2014):

- Restricting tobacco product advertising and marketing to youth
  - Limiting the color and design of packaging and advertisements, including audiovisual advertisements[12]
  - Banning tobacco product sponsorship of sporting or entertainment events under the brand names of cigarettes or smokeless tobacco
  - Banning free samples of cigarettes and brand-name non-tobacco promotional items
- Prohibiting "reduced harm" claims, including "light," "low," or "mild," without an FDA order to allow marketing
  - Requiring industry to submit marketing research documents
- Requiring bigger, more prominent warning labels for cigarettes and smokeless tobacco products
  - Packaging and advertisements for cigarettes and smokeless tobacco must have revised warning labels with a larger font size.

[11] When a manufacturer obtains a marketing order, the manufacturer has complied with the requirements under the FD&C Act to bring its product to market. While FDA may issue a marketing order for a tobacco product to be marketed, the order does not necessarily mean that the tobacco product is safe or "approved" (FDA, 2015b; Miner, 2012).

[12] The implementation of the provision is uncertain due to pending litigation.

Font colors are limited to white on a black background or black on a white background.
 - — Cigarette package health warnings will be required to cover the top 50 percent of both the front and rear panels of the package, and the nine specific warning messages must be equally and randomly displayed and distributed in all areas of the United States. These messages must be accompanied by color graphics showing the negative health consequences of smoking cigarettes.
 - — Smokeless tobacco package warnings must cover 30 percent of the two principal display panels, and the four specific required messages must be equally and randomly displayed and distributed in all areas of the United States.

- Creating an enforcement action plan for advertising and promotion restrictions
  - — FDA published a document titled "Enforcement Action Plan for Promotion and Advertising Restrictions" (FDA, 2010).
  - — The action plan details FDA's current thinking on how it intends to enforce certain requirements under the Tobacco Control Act.

Although not explicitly stated in the Tobacco Control Act, FDA may develop and disseminate public education campaigns that inform the public about the dangers of tobacco products. In February 2014, FDA launched nationally its first youth tobacco prevention campaign, called "The Real Cost," across multiple media platforms, including television, radio, print, and online (FDA, 2015a). The goal of the campaign is to educate at-risk youth aged 12 to 17 who are open to smoking or already experimenting with cigarette use and, by educating them, reduce initiation rates and the prevalence of tobacco use among this population. The campaign will air in more than 200 markets across the country for more than 1 year. In the coming years, FDA plans to develop more youth tobacco prevention campaigns that will target other audiences, including multicultural, rural, and lesbian, gay, bisexual, and transgender youths (Hamburg, 2014).

Sales FDA can also impose restrictions on the retail sales to youth of cigarettes and smokeless tobacco, including (FDA, 2014):

- Requiring proof of age to purchase these tobacco products (the federal minimum age to purchase is 18)
- Requiring face-to-face sales, with certain exemptions for vending machines and self-service displays in adult-only facilities
- Banning the sale of packages of fewer than 20 cigarettes

Limitations FDA does not have the authority to ban certain specified classes of tobacco products,[13] to require the reduction of nicotine yields in tobacco products to zero, to require prescriptions to purchase tobacco products, to reduce the minimum age to purchase tobacco products, or to ban face-to-face tobacco sales in any particular category of retail outlet. The Tobacco Control Act also preserves the authority of state, local, and tribal governments to regulate tobacco products in certain specific respects. It prohibits, with certain exceptions, state and local requirements that are different from, or in addition to, requirements under the provisions of the FD&C Act relating to specified areas.[14]

New tobacco products The Tobacco Control Act defines a tobacco product as any product "made or derived from tobacco" that is not a drug, device, or combination product. In April 2014, FDA proposed to deem all products that meet the definition of a tobacco product to be subject to the FD&C Act, as amended by the Tobacco Control Act (HHS, 2014a). This would either cover "all other categories of products, except accessories of a proposed deemed tobacco product, that meet the statutory definition of 'tobacco product' in the FD&C Act" or else "extend the Agency's 'tobacco product' authorities to all other categories of products, except premium cigars and the accessories of a proposed deemed tobacco product, that meet the statutory definition of 'tobacco product' in the FD&C Act" (HHS, 2014a, p. 23142). The newly covered products would include e-cigarettes, cigars, pipe tobacco, and hookah tobacco. Now that a 105-day public comment period has ended as of August 8, 2014, FDA will review the comments before issuing a final regulation (TCLC, 2014a,b). If finalized in its current form, the deeming rule will give FDA the authority to restrict the sale of newly covered tobacco products to minors below the age of 18, prohibit their sales in vending machines except in adults-only venues, prohibit free samples, and require a health warning on package labels (TCLC, 2014b). However, FDA would not automatically claim the authority to restrict the marketing of newly covered products (except false or misleading advertising) or prohibit the use of flavorings such as in e-cigarettes, even though it can do so for traditional cigarettes. FDA would retain the authority to take these actions in the future (TCLC, 2014b).

[13] FDA cannot ban all cigarettes, all smokeless tobacco products, all little cigars, all cigars other than little cigars, all pipe tobacco, or all roll-your-own tobacco products.

[14] Federal Food, Drug, and Cosmetic Act of 1938, Chapter IX. Public Law 75-717, 75th Cong. (1938).

Roles of Other Federal Agencies for Tobacco

Other federal agencies are involved in tobacco regulation in various ways, including prevention, enforcement, cessation, and compliance (Leischow et al., 2010). The Center for Drug Evaluation and Research, a department within FDA, plays a significant role in the regulation of smoking cessation medications. Other agencies within HHS with tobacco-related responsibilities include CDC;[15] the Substance Abuse and Mental Health Services Administration;[16] the National Institutes of Health, including the National Cancer Institute[17] and the National Institute on Drug Abuse;[18] the Health Resources and Services Administration;[19] the Centers for Medicare & Medicaid Services;[20] and the Indian Health Service.[21] Within HHS, the departments that have tobacco-related responsibilities have regular meetings to foster communication across the department (Leischow et al., 2010). Other agencies that have tobacco-related responsibilities include the Alcohol and Tobacco Tax and Trade Bureau;[22] the Bureau of Alcohol, Tobacco,

[15] "CDC, through the Office on Smoking and Health (OSH), is the lead federal agency for comprehensive tobacco prevention and control. OSH is a division within the National Center for Chronic Disease Prevention and Health Promotion, which is located within CDC's Coordinating Center for Health Promotion. Originally established in 1965 as the National Clearinghouse for Smoking and Health, OSH is dedicated to reducing the death and disease caused by tobacco use and exposure to secondhand smoke" (CDC, 2014b).

[16] The Substance Abuse and Mental Health Services Administration oversees implementation of the Synar Amendment, which requires states to have laws in place prohibiting the sale and distribution of tobacco products to minors, and the enforcement of those laws.

[17] The National Cancer Institute's Tobacco Control Research Branch leads and collaborates on research and disseminates evidence-based findings to prevent, treat, and control tobacco use. Additionally, in partnership with Legacy, the National Cancer Institute created the Tobacco Research Network on Disparities to help facilitate the elimination of health disparities related to tobacco.

[18] The National Institute on Drug Abuse works with FDA and supports a wide variety of research on tobacco from basic science to tobacco control policy.

[19] The Health Resources and Services Administration aims to have 100 percent of its health center grantees adopt formal tobacco prevention and cessation programs.

[20] The Centers for Medicare & Medicaid Services cover treatment of tobacco-related illness and cessation counseling in certain circumstances.

[21] The Tobacco Control and Prevention Program of the Indian Health Service seeks to improve the physical, mental, social, and spiritual health of American Indians and Alaska Natives through the prevention and reduction of tobacco-related disease.

[22] The Alcohol and Tobacco Tax and Trade Bureau (under the U.S. Department of Treasury) assures compliance with federal tobacco permitting and collects federal tobacco taxes.

Firearms and Explosives;[23] the Federal Trade Commission;[24] and the U.S. Environmental Protection Agency.[25]

State, Local, and Tribal Authority

As noted earlier, the Tobacco Control Act preserves the authority of state, local, and tribal governments to regulate tobacco products in certain specific respects. Under the new law, state and local governments can engage in a large array of tobacco policies aimed at reducing tobacco use and improving health in addition to authorities they had before the law was passed (such as taxation). They can use communication interventions to convey the risks and harms of tobacco through many avenues, including print, other media, and the Internet. They can also engage in and increase access to cessation programs aimed at helping tobacco users stop. Other levers often used by states and localities include raising tobacco taxes (which range from a low of 17 cents per pack in Missouri to a high of $4.35 per pack in New York) (Henchman and Drenkard, 2014); passing smoke-free laws that apply to restaurants, bars, and workplaces; restricting the sale, distribution, and possession of tobacco products; and implementing tax evasion and anti-smuggling measures (CDC, 2014a; TCLC, 2009). The Tobacco Control Act permits state and local governments to:

- Expand the current requirements of the Tobacco Control Act that limit advertisements for cigarettes and smokeless tobacco to black-and-white text to apply to advertisements for cigars and other tobacco products as well
- Prohibit the display of "power walls" of cigarette packages at retail outlets
- Limit the number and size of tobacco advertisements at retail outlets
- Require that tobacco products (and advertisements) be kept a minimum distance from cash registers

States and localities can also impose minimum age and other sale restrictions, retail density laws, fire-safe laws, reporting requirements (such as ingredients), and point-of-sale warnings, among others. However, they can-

[23] Under the U.S. Department of Justice, the Bureau of Alcohol, Tobacco, Firearms and Explosives aims to "reduce alcohol smuggling and contraband cigarette trafficking activity, divest criminal and terrorist organizations of monies derived from this illicit activity and significantly reduce tax revenue losses to the States" (ATF, 2015).

[24] The Federal Trade Commission investigates unfair tobacco industry business practices and advertisements and enforces laws that address these practices.

[25] The U.S. Environmental Protection Agency regulates pesticides used on tobacco plants.

not place requirements on cigarette or smokeless tobacco product labeling or on the content of cigarette or smokeless tobacco advertisements; those are under the jurisdiction of FDA.

Finally, the nations of Indian Country, as sovereign entities, have significant regulatory powers over tobacco. Multiple tribes now produce, market, and sell tobacco products and view the manufacturing and sales of tax-free tobacco products as a revenue opportunity, a benefit to tribal economic development, and, perhaps most importantly, an exercised right of their sovereign statuses. State excise taxes do not apply to cigarettes sold to tribal members on tribal land (Samuel et al., 2012). Furthermore, even though federal law requires the collection and remittance of excise taxes of cigarette sales to non-tribal members, states cannot force tribes to collect them (Samuel et al., 2012). Thus, cigarettes sold on reservations may not include any state excise taxes, resulting in significantly lower costs for consumers (Hyland et al., 2005). One result of these low costs is that individuals or sellers on the black market skirt taxes by purchasing cigarettes on reservations (Kurti et al., 2012). It is important to note that although states have no authority over tribal nations, many states and tribes have entered into compact agreements regarding taxation.[26] The federal government has regulatory authority over tribal tobacco products, but the extent of this authority is still somewhat unclear.

TOBACCO USE BEHAVIOR

To use computational modeling effectively to examine the impact of policy or interventions on tobacco use, modelers need to understand not only the tobacco environment but also tobacco use behavior by individuals. This section provides a high-level overview of some of the theoretical and empirical concepts about initiation and cessation of tobacco use among youth and adults and discusses characteristics of tobacco products, particularly their addictive nature, that are important to consider when developing a computational model. The onset, progression, and cessation of tobacco use among youth and adults are complex and multifactor processes, and they have been conceptualized from the perspective of many different fields, including brain disease, genetics, economics, psychology, and sociology. Among the various perspectives, this section focuses primarily on key drivers from the social and behavioral sciences because there is a large body of literature looking at tobacco-related behavior from these perspectives. It deserves mention, however, that perspectives from other fields have been responsible for major contributions to understanding tobacco use behavior, and these need

[26] For more information on tribal tax codes and agreements, see NCAI, 2015, and https://www.sos.ok.gov/gov/tribal.aspx.

to be taken into account when attempting to predict responses to policies. The perspectives used in this chapter will vary, depending on the particular question being addressed.

Tobacco Use—Factors to Be Considered When Developing Models

Tobacco use is a complex behavior that is determined by a wide range of factors that need to be considered in assessing a policy's effects on the prevalence of tobacco use and its health consequences. Tobacco use initiation and tobacco use cessation are understood to be distinct multistep processes that are influenced by different but sometimes overlapping factors. For example, the reasons that people initiate the use of tobacco products (e.g., social influence) are often different from their reasons for continuing to use the products (e.g., addiction). The goal of this section is to summarize the current understanding of some of the factors leading to behavior change that policy makers need to consider when predicting or assessing the impact of tobacco control interventions. Computational models designed to forecast the effects of tobacco policies need to take these factors into account early in the development process as part of the conceptual framework.

Biological, psychological, social–contextual, and economic factors, among others, contribute to the development, maintenance, and change of health behavior patterns such as tobacco use. Conceptual frameworks and theories offer systematic ways of understanding key determinants of behaviors and guide the search to identify the data and information that are needed to predict behavior. For example, cognitive–behavioral models (Glanz et al., 2005)[27] are drawn from the social and behavioral sciences. Examples include the health belief model[28] (Hochbaum, 1958) and social cognitive theory[29] (Bandura, 1986), which have been used to predict tobacco use behavior at the intrapersonal levels and interpersonal levels, respectively. Empirical testing of these theories has established that the constructs included in these models influence initiation (HHS, 2014b)

[27] Cognitive–behavioral models are based on the assumptions that behavior is mediated by cognitions (i.e., what people know and think affects how they act); knowledge is necessary for, but not sufficient to produce, most behavior changes; and perceptions, motivations, skills, and the social environment are key influences on behavior (Glanz et al., 2005).

[28] The theory suggests that health behavior is determined by personal beliefs or perceptions about a disease and the approaches available to decrease its occurrence. Four perceptions serve as the main constructs of the health belief model: perceived seriousness, perceived susceptibility, perceived benefits, and perceived barriers. Each of these perceptions, individually or in combination, can be used to explain health behavior. The model has added other constructs, including cues to action, motivating factors, and self-efficacy (Hayden, 2014).

[29] This theory posits that health is a function of factors that exist across intrapersonal, interpersonal, and community levels.

and cessation of tobacco use (Prochaska et al., 2008). Empirical research has also established a number of biological mechanisms relevant to tobacco use behavior at the intrapersonal level, particularly tobacco use cessation. For example, genetic susceptibility to nicotine and the physiologic pathways specific to serotonin and dopamine receptors and processing in the brain influence an individual's ability to quit tobacco use (HHS, 2010). Finally, factors such as time preference and discount rates (i.e., how individuals value costs and benefits that occur in the future versus those in the present), risk aversion, price, marketing, the development of information on the harms of smoking, provision of information by the government, social networks, and behavioral economics concepts such as heuristics all are considered in economic models and empirical analyses aimed at understanding tobacco use behavior (Cawley and Ruhm, 2011; Chaloupka, 1991; Chaloupka and Warner, 2000; Smith et al., 2014).

Although a variety of conceptual frameworks and theories have been used in efforts to understand tobacco use behavior, many of them have emphasized an ecological perspective, which asserts that an individual's behavior both shapes and is shaped by the social environment (Bandura, 1986; McLeroy et al., 1988; Sallis et al., 2008). In other words, an individual's behavior is understood to affect and be affected not only by intrapersonal characteristics such as knowledge, attitudes and beliefs, but also by interpersonal factors like peer influence and community-level factors like social norms and policies (see Figure 2-3).

Concept	***Definition***
Intrapersonal Level	Individual characteristics that influence behavior, such as knowledge, attitudes, beliefs, and personality traits
Interpersonal Level	Interpersonal processes and primary groups, including family, friends, and peers that provide social identity, support, and role definition
Community Level	
Institutional Factors	Rules, regulations, policies, and informal structures, which may constrain or promote recommended behaviors
Community Factors	Social networks and norms, or standards, which exist as formal or informal among individuals, groups, and organizations
Public Policy	Local, state, and federal policies and laws that regulate or support healthy actions and practices for disease prevention, early detection, control, and management

FIGURE 2-3 An ecological perspective: Levels of influence.
SOURCE: Glanz et al., 2005.

Moreover, the factors influencing tobacco use at the intrapersonal, the interpersonal, and the community levels interact in a complex web, with the factors both acting directly on tobacco use behaviors and also altering each other's impact on tobacco use behavior.[30] The next section uses an ecological perspective to highlight the salient factors that drive initiation and cessation of tobacco use—and, specifically, of cigarette smoking, because most studies to date have focused on cigarette smoking. The ecological perspective is often used in public health to frame complex behavior, and so it is used in the following section to guide the discussion on tobacco use initiation and cessation. However, it is important to note that there are other perspectives in the research literature that can be valuable when studying tobacco use behavior, such as those from biology or economics, and perspectives from these fields are included where appropriate.

Smoking Initiation

The onset and progression of tobacco use among young people, from adolescence into young adulthood, is a dynamic, multistage process influenced by multiple determinants. As shown in Figure 2-3, these determinants operate at intrapersonal, interpersonal, and other (e.g., social, community, public policy) levels. Adolescence and young adulthood are critical periods in the life course when tobacco use may be especially appealing and even functional, as substantial research indicates that it is not merely a "rational choice" in these developmental phases. Developmentally speaking, young people progress from experimentation with tobacco use, to intermittent use, to regular use and dependence. Not all young people progress through all stages, and movement across these stages can be both forward and backward. Nearly 90 percent of adult daily smokers report that they started using cigarettes before the age of 18, and two-thirds made the transition to daily use during adolescence (HHS, 2012). Longitudinal studies show that it takes 3 years on average to move from experimentation to regular (i.e., daily) use, with considerable variation between individuals in both the process and the timing (HHS, 2012; Mayhew et al., 2000). Before and during each of these stages, attitudes and beliefs about the utility of tobacco use are formed that can drive movement forward or backward between stages. Identifying key factors that drive continued use or that interrupt progress

[30] One variable may mediate the effect (i.e., be on the causal pathway) of a second variable's influence on tobacco use initiation or cessation. For example, social norms about tobacco use in a particular school can alter peer influence in that school, which in turn will affect the initiation of tobacco use among the youth in that school. Alternately, variables can moderate (i.e., change) the strength of the relationship between a second variable and tobacco use. For example, the impact of peer influence on tobacco use initiation among youth may be moderated by the effects of parent tobacco use at home.

along this continuum will be critical to any modeling process designed to predict tobacco use behaviors in young populations.

At the intrapersonal level, beliefs about the health and social consequences of tobacco use, decision-making capabilities, and the ability to regulate one's behavior (i.e., risk taking) all help predict the onset and progression of tobacco use among young people. In addition to these cognitive processes, implicit attitudes (e.g., liking smoking or being willing to date a tobacco user) are also related to tobacco use among youth (HHS, 2012). Behavioral factors such as poor academic performance are also correlates of initiation and continued tobacco use by youth. Differences in tobacco use behaviors among youths with different levels of academic success persist and grow as the youths move from adolescence into young adulthood, so that by the time they reach young adulthood, the prevalence of tobacco use among non-college-going youth is twice that among those attending college (HHS, 2012).

Many researchers and public health experts believe that one of the most important factors affecting youth tobacco use is social influence, which occur at the interpersonal level (HHS, 2012). Peer influences are especially salient and strong. For example, young people often overestimate the prevalence of tobacco use among their peers (i.e., people of the same age), which is particularly important because perceptions that one's peers use tobacco consistently predict an individual's tobacco use. Both selection (i.e., choosing new friends who use tobacco) and socialization (i.e., being influenced by existing friends who use tobacco) are relevant to the movement between stages of tobacco use described earlier (HHS, 2012). Furthermore, research shows that tobacco use behaviors among and within these peer networks are influenced by group-level norms (e.g., school norms) and attempts to be liked by others in the group (HHS, 2012). As cigarette smoking has become less normative in the United States, recent research suggests that youth who self-identify as belonging to deviant peer groups are most likely to be smokers (HHS, 2012).

At the interpersonal level, family influences on youth tobacco use behavior can also be strong. The research on parental influence, including parental disapproval of tobacco use, parent tobacco use behaviors, and parenting practices (e.g., monitoring a child's tobacco use, even as a young adult) suggests that these factors often moderate the influence of other factors, such as peer influence (HHS, 2012). The use of tobacco products among older siblings is also a predictive factor in youth tobacco use (HHS, 2012).

The intrapersonal and interpersonal factors listed above all interact in a complex web that can vary among individuals. These factors have stronger, more direct, and more immediate effects on youth tobacco use than other macro-level factors such as school climate and community norms about tobacco use (HHS, 2012). However, macro-level factors must not be discounted, as they are the context in which these influences take shape. This context, in turn, is influenced by macro-level interventions, such as various tobacco policies (e.g., increased taxes on tobacco products) and communication campaigns (e.g., social marketing).

Smoking Cessation

Smoking cessation, like smoking initiation, is conceptualized as a multistep process. To succeed in stopping tobacco use, an individual must first decide to make an attempt to quit and then succeed in that attempt. Most individual quit attempts, even those made using the most effective contemporary treatments, do not succeed (Hatsukami et al., 2008). Instead, the tobacco user relapses (that is, returns to smoking), most often within the first week (Hughes et al., 2004). After 3 months, the likelihood of resuming tobacco use decreases, and a quit attempt that lasts for 6 or 12 months is generally considered to represent long-term successful cessation of tobacco use (HHS, 2010). However, many tobacco smokers return to smoking even after 12 months of abstinence (HHS, 2010). This multistep process is considered to be influenced primarily by internal factors, both biological and psychosocial, although interpersonal and macro factors such as social support are also important.

At the intrapersonal level, biological factors, specifically physiologic dependence on nicotine, influence an individual's ability to succeed at quitting smoking. Nicotine is the major chemical component of tobacco smoke responsible for causing physiologic dependence on cigarettes. An individual's risk of nicotine addiction depends on the dose of nicotine delivered and the way it is delivered (HHS, 2010). There is also evidence that some people are more predisposed to becoming addicted because of psychological or genetic factors or both (HHS, 2012). When an individual inhales cigarette smoke, nicotine is rapidly delivered to the brain, binding to nicotinic cholinergic receptors and activating the release of dopamine and other neurotransmitters that reinforce smoking and the behaviors associated with smoking (Benowitz, 2010). As an individual's cigarette smoke exposure increases, the number of nicotine receptors increases, producing tolerance to higher doses of nicotine (Benowitz, 2008b; HHS, 2010). At that point, when nicotine levels decrease, an individual may experience withdrawal symptoms, including irritability, impatience, difficulty concentrating, an anxious or depressed mood, and an increased appetite. Smoking a cigarette alleviates these unpleasant symptoms. These symptoms begin within a few hours of smoking cessation, peak at 48 to 72 hours, and gradually diminish over weeks, although the duration and severity of nicotine withdrawal depend on the degree of nicotine addiction (HHS, 2010). This explains why smokers often relapse (return to smoking) in the first hours and days after stopping smoking. To stop smoking, an addicted individual must manage and overcome the symptoms of nicotine withdrawal.

In addition to the addiction to nicotine, tobacco smoking is maintained by various behavioral factors, especially related habitual behaviors. For example, after repeated pairings of smoking with the end of a meal, a smoker comes to associate smoking with finishing a meal (HHS, 2010). Thereafter, finishing a meal triggers an urge to smoke. Additionally, smoking appears to increase an individual's enjoyment of other reinforcers. For example, many individuals say that they crave cigarettes when drinking alcohol (HHS, 2010). It appears that nicotine has the effect of enhancing an individual's pleasure from the other reinforcer (in this case, alcohol). The dual challenge of overcoming not only nicotine withdrawal symptoms but also learned behavioral associations with smoking causes many individual quit attempts to fail.

Various other factors, from intrapersonal to macro-level factors, can also influence the process of smoking cessation. For example, individuals may be motivated to stop smoking if they believe that the benefits of quitting outweigh the pleasures of continuing to smoke. The decision is a balance of the perceived threat of continuing to smoke and the beliefs of individuals that stopping smoking will benefit them. Another factor is a smokers' confidence in his or her ability to succeed at quitting (i.e., self-efficacy). In many different treatment trials this has been shown to be an independent predictor of cessation success (Ockene et al., 2000). Finally, macro factors, such as social support from family and friends, cohabitation with smokers, and a tobacco user's access or adherence to treatment can influence cessation through the intrapersonal and intrapersonal levels (TUDGP, 2008).

Finally, it is important to note that socio-demographic factors (e.g., age, gender/sex, race/ethnicity, socioeconomic status) often moderate the impact of intrapersonal, interpersonal, and macro-level determinants described above and the speed with which individuals progress through these stages of tobacco use. One caveat is that the majority of empirical research to date has been specific to cigarette smoking, so that there is limited etiologic research available concerning the developmental processes and pathways of initiation for other tobacco products. Therefore, in modeling exercises one must understand that the factors important to the onset and progression of cigarette smoking for a white, affluent, 12-year-old girl may be very different from the factors important to the progression of cigar smoking for a 15-year-old African American boy. Both the factors themselves and the magnitude of the impact that the factors have on the onset, progression, and cessation of tobacco use among youth and young adults can be different.

WHY USE AGENT-BASED MODELS TO EXPLORE TOBACCO USE?

Existing models used in tobacco control have focused mostly on determining the long-term dynamics of population-level tobacco rates, either by extrapolating the status quo into the future or by projecting the consequences of policy interventions. For example, these models have been used to project changes in smoking prevalence and associated morbidity and mortality in the U.S. population over the next few decades, assuming that initiation and cessation rates remain fixed at current levels (HHS, 2014b). Those figures have then been contrasted with the morbidity and mortality implied by the same models when initiation and cessation rates are affected by various policy interventions, using what is known from past experience concerning how such interventions have affected the initiation and cessation rates.

These analyses have employed almost exclusively aggregate compartmental/system dynamics models, which assume a large degree of homogeneity among the population. In other words, these models assume that members of the population can be classified according to a limited number of distinguishing characteristics (for example, never, current, or former smokers) and that within each group in that classification the individuals behave identically. Additionally, the majority of these models do not consider social interactions among members of the population, and when they do, they assume that individuals within their unique groups are perfectly mixed and thus have the same chance of interacting with each other.

In general, this work has also treated certain important smoking processes as exogenous to the models (that is, as coming from outside a model and thus unexplained by the model). Tobacco use initiation, cessation, and relapse have been specified as externally supplied probabilities (that is, probabilities that are not explained by the model) that affect individuals at certain periods of their lives. Although these models do allow for policy interventions to affect the various probabilities, the models provide no details about the underlying mechanisms that generate such chances and their potential feedback relationship with the tobacco use rates they generate.

These aggregate population models have been very useful in determining the overall magnitude of the tobacco epidemic and its likely trajectory; however, given the increased complexity of the tobacco use landscape, it is becoming evident that policy makers need to better understand and model explicitly the essential social- and individual-level processes of tobacco use behavior (e.g., the mechanisms of initiation, cessation, and relapse) in order to anticipate as accurately as possible the effects of policy interventions. There is evidence that processes such as tobacco use initiation, cessation, and relapse are at least partially driven by social interactions. For example, the presence and strength of connection to friends or parents who smoke is likely to have an impact on a person's decision to initiate the smoking behavior, particularly among youth and young adults (HHS, 2012). Factors related to social interactions also play a role in smoking among adults, although there is insufficient evidence to suggest that these factors are paramount as compared to smoking initiation among youth and young adults. Similarly, the process of quitting smoking is influenced by interactions with other individuals (Chandola et al., 2004; Herd et al., 2009; Hitchman et al., 2014; Hymowitz et al., 1997; Westmaas et al., 2010). There is also ample evidence that living with individuals who smoke is a strong predictor of relapse among those who have recently quit (Garvey et al., 1992; Mermelstein et al., 1986). (See Chapter 3 for an in-depth discussion on social interactions.)

Analysis of survey data can help researchers identify the nature and strength of these social determinants of smoking behavior at the individual level, but computational models in general—and ABMs specifically—are needed to estimate the total population effects of those individual interactions. (For a description of ABMs see Chapter 1 and Appendix A.) These models can account for individuals' differences and the multiple ways in which such individuals are influenced and can influence each other in order to estimate the combined effect of the multiple processes that constitute tobacco use behavior. These models can also account for important feedback mechanisms that have been, for the most part, ignored by existing aggregate models. For example, if a peer effect on smoking initiation is considered, a decline in smoking rates among the population would translate into a decline in smoking initiation, producing a cascade effect that would drive down smoking prevalence faster than what has been anticipated by traditional models. Although it is not guaranteed that ABMs will answer all policy questions, they may be able to inform those policies with underlying behavioral questions more fully than other modeling methods. Specifically, they are likely to inform the development, implementation, and enforcement of policies that are intended to influence behaviors. It is important to note that ABMs that do not focus on individual tobacco use behavior may also be useful to FDA or other tobacco control policy scientists. Examples include models of how the development of new tobacco products disrupt existing industry and retailer practices or models of community-level policies at the point of sale that are designed to affect retailer behavior (e.g., advertising).

In sum, given the strong social component inherent in tobacco use onset, cessation, and relapse and the heterogeneity of the relevant social interactions, ABMs have the potential to be an essential tool in assessing the effects on policies to control tobacco. These models could clarify the net effects that enacted policies have on a complex social environment and potentially inform inputs for aggregate population models, which focus on the long-term consequences of such policies. Additionally, ABMs that explicitly model critical processes of tobacco use, such as initiation, cessation, and relapse, could be used to help answer many larger policy-relevant questions faced by FDA. Specific questions in the tobacco control field that researchers might want to model include what is the public health impact of lowering the nicotine content of combustible cigarettes to non-addictive levels? What is the impact of banning flavorings in electronic cigarettes? What are the impacts of competing media or education campaigns on reducing tobacco use? All of these questions require an understanding of the underlying behavioral mechanisms involved in initiation and cessation, and answering them would require a specific model of those processes.

Within the modeling community, it is often said that models need to be motivated by a specific question to be effective (Bankes, 1993; Bruch and Atwell, 2013; CREM, 2009; Macal and North; NRC, 2007, 2014). This may often be the case, as illustrated above, but the processes or mechanisms underlying these questions—what happens in the black box between policy implementation and potential behavior change—often need to be the focal point of the model before the specific question is addressed.

> Finding 2-1: The committee finds that for many tobacco control policy questions, several key underlying processes—initiation, cessation, and relapse, among others—drive overall rates of tobacco use and have a strong social interaction component. An agent-based model could be a useful tool to represent these processes.

In the case of tobacco, a useful path will be to develop models of these processes first and then to apply them to the specific policy question. This does not imply that all efforts should be put into a single model of social processes that would then be applied to many different questions. Rather, accurately representing the underlying process of initiation, cessation, and relapse is, in some cases, essential to the development of a model of tobacco use behavior. It is important to note that there are other features of tobacco control policy that are not directly related to initiation and cessation (e.g., tobacco companies responding to FDA regulatory changes in an attempt to undermine those changes), so the modeling decision to focus on a specific policy question versus initiation or cessation needs to discussed early in model conceptualization.

This section outlines the conditions in which an ABM could be useful to inform tobacco control policy. It is difficult to identify specific domains that all policy-relevant ABMs would need to incorporate. Given the large range of factors and agents in the tobacco regulatory environment (as illustrated in this chapter), ABMs developed by CTP or others will require consultation before development begins (and throughout the lifespan of the model) with relevant stakeholders, subject-matter experts, end users (e.g., decision makers), and the modeler to decide what domains and ABM characteristics would be appropriate for the intended purpose of the model (see Chapter 4). Although this report does not identify specific attributes or domains a tobacco control ABM would require—as that would vary depending on the purpose of the ABM as each policy question involves different agents and levels of interaction—it offers advice on the conditions in which ABM could be appropriate, important data considerations, and an evaluation framework that CTP can use in their future development of ABMs.

CONCLUDING COMMENTS

The description of the tobacco environment provided in this chapter outlines the difficult, but necessary, task of using models for tobacco use. Given that the results of tobacco control models could be used to inform real-time decisions by policy makers, it is critical to ensure that the modeling methods used are suitable for the question at hand and that they provide results that are reliable. In the next chapter, the committee offers guidance on using individual-level models to aid policy decisions.

REFERENCES

Abrams, D. B. 2014. Promise and peril of e-cigarettes: Can disruptive technology make cigarettes obsolete? *JAMA* 311(2):135–136.

Agaku, I. T., B. A. King, C. G. Husten, R. Bunnell, B. K. Ambrose, S. S. Hu, E. Holder-Hayes, H. R. Day, and Centers for Disease Control and Prevention. 2014. Tobacco product use among adults—United States, 2012–2013. *Morbidity and Mortality Weekly Report* 63(25):542–547.

Alamar, B. C., and S. A. Glantz. 2004. The tobacco industry's use of Wall Street analysts in shaping policy. *Tobacco Control* 13(3):223–227.

APSC (Australian Public Service Commission). 2007. *Tackling wicked problems: A public policy perspective*. Australian Government, Australian Public Service Commission. http://www.apsc.gov.au/__data/assets/pdf_file/0005/6386/wickedproblems.pdf (accessed April 1, 2015).

ATF (Bureau of Alcohol, Tobacco, Firearms and Explosives). 2015. *Alcohol & tobacco*. https://www.atf.gov/content/alcohol-and-tobacco (accessed February 18, 2015).

Bandura, A. 1986. *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Bankes, S. 1993. Exploratory modeling for policy analysis. *Operations Research* 41(3):435–449.

Bauld, L., K. Angus, and M. de Andrade. 2014. *E-cigarette uptake and marketing: A report commissioned by Public Health England*. Institute for Social Marketing, University of Stirling, UK Centre for Tobacco & Alcohol Studies, Public Health England.

Benowitz, N. L. 2008a. Clinical pharmacology of nicotine: Implications for understanding, preventing, and treating tobacco addiction. *Clinical Pharmacology and Therapeutics* 83(4):531–541.

———. 2008b. Neurobiology of nicotine addiction: Implications for smoking cessation treatment. *The American Journal of Medicine* 121(4 Suppl 1):S3–S10.

———. 2010. Nicotine addiction. *The New England Journal of Medicine* 362(24):2295–2303.

Bhatnagar, A., L. P. Whitsel, K. M. Ribisl, C. Bullen, F. Chaloupka, M. R. Piano, R. M. Robertson, T. McAuley, D. Goff, N. Benowitz, American Heart Association Advocacy Coordinating Committee, Council on Cardiovascular and Stroke Nursing, Council on Clinical Cardiology, and Council on Quality of Care and Outcomes Research. 2014. Electronic cigarettes: A policy statement from the American Heart Association. *Circulation* 130(16):1418–1436.

Blosnich, J., J. G. Lee, and K. Horn. 2013. A systematic review of the aetiology of tobacco disparities for sexual minorities. *Tobacco Control* 22(2):66–73.

Bruch, E., and J. Atwell. 2013. Agent-based models in empirical social research. *Sociological Methods & Research* 1–36.

Cawley, J., and C. Ruhm. 2011. The economics of risky health behaviors. In *Handbook of health economics*. Vol. 2. Cambridge, MA: National Bureau of Economic Research. Pp. 95–199.

CDC (Centers for Disease Control and Prevention). 2014a. *Best practices for comprehensive tobacco control programs*. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

———. 2014b. *Smoking and tobacco use: About this office*. http://www.cdc.gov/tobacco/about (accessed February 19, 2015).

Chaloupka, F. J. 1991. Rational addictive behavior and cigarette smoking. *Journal of Political Economy* 99(4):722–742.

Chaloupka, F. J., and K. E. Warner. 2000. The economics of smoking. In *Handbook of health economics edition 1*. Vol. 1, edited by A. J. Culyer and J. P. Newhouse. Amsterdam, Netherlands: North Holland, Elsevier. Pp. 1539–1627.

Chandola, T., J. Head, and M. Bartley. 2004. Socio-demographic predictors of quitting smoking: How important are household factors? *Addiction* 99(6):770–777.

CREM (Council for Regulatory Environmental Modeling). 2009. *Guidance on the development, evaluation, and application of environmental models*. Washington, DC: Office of the Science Advisor, Council for Regulatory Environmental Modeling, U.S. Environmental Protection Agency.

Crosse, M. 2014. *Tobacco products: FDA spending and new product review time frames (GAO-14-508T)*. Washington, DC: U.S. Government Accountability Office. http://www.gao.gov/assets/670/662360.pdf (accessed April 1, 2015).

Dilley, J. A., C. Spigner, M. J. Boysun, C. W. Dent, and B. A. Pizacani. 2008. Does tobacco industry marketing excessively impact lesbian, gay and bisexual communities? *Tobacco Control* 17(6):385–390.

Dorfman, L., and L. Wallack. 1993. Advertising health: The case for counter-ads. *Public Health Reports* 108(6):716–726.

Fairchild, A. L., R. Bayer, and J. Colgrove. 2014. The renormalization of smoking? E-cigarettes and the tobacco "endgame." *New England Journal of Medicine* 370(4):293–295.

Fallin, A., A. Goodin, Y. O. Lee, and K. Bennett. 2014. Smoking characteristics among lesbian, gay, and bisexual adults. *Preventive Medicine* 73:123–130.

FDA (U.S. Food and Drug Administration). 2009. *Tobacco product fees*. http://www.fda.gov/ForIndustry/UserFees/TobaccoProductFees/default.htm (accessed January 22, 2015).

———. 2010. *Enforcement action plan for promotion and advertising restriction*. http://www.fda.gov/TobaccoProducts/GuidanceComplianceRegulatoryInformation/ucm271744.htm (accessed February 25, 2015).

———. 2014. *Overview of the Family Smoking Prevention and Tobacco Control Act*. http://www.fda.gov/downloads/TobaccoProducts/GuidanceComplianceRegulatoryInformation/UCM336940.pdf (accessed January 22, 2015).

———. 2015a. *The real cost: Overview*. http://www.fda.gov/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/AbouttheCenterforTobaccoProducts/PublicEducationCampaigns/TheRealCostCampaign/ucm388656.htm (accessed February 25, 2015).

———. 2015b. *Tobacco product marketing orders*. http://www.fda.gov/TobaccoProducts/Labeling/MarketingandAdvertising/ucm339928.htm (accessed January 22, 2015).

Fu, S. S., D. Burgess, M. van Ryn, D. K. Hatsukami, J. Solomon, and A. M. Joseph. 2007. Views on smoking cessation methods in ethnic minority communities: A qualitative investigation. *Preventive Medicine* 44(3):235–240.

Garvey, A. J., R. E. Bliss, J. L. Hitchcock, J. W. Heinold, and B. Rosner. 1992. Predictors of smoking relapse among self-quitters: A report from the normative aging study. *Addictive Behaviors* 17(4):367–377.

Gfroerer, J., S. R. Dube, B. A. King, B. E. Garrett, S. Babb, and T. McAfee. 2013. Vital signs: Current cigarette smoking among adults aged greater than or equal to 18 years with mental illness—United States 2009–2011. *Morbidity and Mortality Weekly Report* 62(5):81–87.

Glanz, K., B. K. Rimer, and NCI. 2005. *Theory at a glance: A guide for health promotion practice*. 2nd ed. NIH Publication No. 05-3896. Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health, National Cancer Institute.

Grana, R., N. Benowitz, and S. A. Glantz. 2014. E-cigarettes: A scientific review. *Circulation* 129(19):1972–1986.

Hajek, P. 2013. Electronic cigarettes for smoking cessation. *Lancet* 382(9905):1614–1616.

Hamburg, M. A. 2014. *FDA public education campaign aims to prevent and reduce youth tobacco use*. http://blogs.fda.gov/fdavoice/index.php/2014/02/fda-public-education-campaign-aims-to-prevent-and-reduce-youth-tobacco-use (accessed November 13, 2014).

Hatsukami, D. K., L. F. Stead, and P. C. Gupta. 2008. Tobacco addiction. *Lancet* 371(9629):2027–2038.

Hatzenbuehler, M. L., N. F. Wieringa, and K. M. Keyes. 2011. Community-level determinants of tobacco use disparities in lesbian, gay, and bisexual youth: Results from a population-based study. *Archives of Pediatrics and Adolescent Medicine* 165(6):527–532.

Hayden, J. 2014. Chapter 4: Health belief model. In *Introduction to health behavior theory*. 2nd ed. Burlington, MA: Jones & Bartlett Learning.

Henchman, J., and S. Drenkard. 2014. Cigarette taxes and cigarette smuggling by state. *Tax Foundation Fiscal Fact*. http://taxfoundation.org/sites/taxfoundation.org/files/docs/FF421.pdf (accessed November 13, 2014).

Herd, N., R. Borland, and A. Hyland. 2009. Predictors of smoking relapse by duration of abstinence: Findings from the International Tobacco Control (ITC) four country survey. *Addiction* 104(12):2088–2099.

HHS (U.S. Department of Health and Human Services). 2010. *How tobacco smoke causes disease: The biology and behavioral basis for smoking-attributable disease: A report of the Surgeon General.* Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention, Office on Smoking and Health.

———. 2012. *Preventing tobacco use among youth and young adults: A report of the Surgeon General* Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

———. 2014a. 21 CFR parts 1100, 1140, 1143 proposed rule. *Federal Register* 79(80):23142–23207.

———. 2014b. *The health consequences of smoking—50 years of progress: A report of the Surgeon General.* Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

Hitchman, S. C., G. T. Fong, M. P. Zanna, J. F. Thrasher, and F. L. Laux. 2014. The relation between number of smoking friends, and quit intentions, attempts, and success: Findings from the International Tobacco Control (ITC) four country survey. *Psychology of Addictive Behaviors*. 28(4):1144–1152.

Hochbaum, G. M. 1958. *Public participation in medical screening programs; a socio-psychological study*. Washington, DC: U.S. Department of Health, Education, and Welfare, Public Health Service, Bureau of State Services, Division of Special Health Services, Tuberculosis Program.

Hughes, J. R., J. Keely, and S. Naud. 2004. Shape of the relapse curve and long-term abstinence among untreated smokers. *Addiction* 99(1):29–38.

Hyland, A., C. Higbee, Q. Li, J. E. Bauer, G. A. Giovino, T. Alford, and K. M. Cummings. 2005. Access to low-taxed cigarettes deters smoking cessation attempts. *American Journal of Public Health* 95(6):994–995.

Hymowitz, N., K. M. Cummings, A. Hyland, W. R. Lynn, T. F. Pechacek, and T. D. Hartwell. 1997. Predictors of smoking cessation in a cohort of adult smokers followed for five years. *Tobacco Control* 6(Suppl 2):S57–S62.

Jamal, A., I. T. Agaku, E. O'Connor, B. A. King, J. B. Kenemer, and L. Neff. 2014. Current cigarette smoking among adults—United States, 2005–2013. *Morbidity and Mortality Weekly Report* 63(47):1108–1112.

King, B. A., S. R. Dube, and M. A. Tynan. 2012. Current tobacco use among adults in the United States: Findings from the National Adult Tobacco Survey. *American Journal of Public Health* 102(11):e93–e100.

Kirkwood, C. 1998. Chapter 1: System behavior and causal loop diagrams. In *System dynamics methods: A quick introduction*. http://www.nutritionmodels.com/copyrighted_papers/Kirkwood1998.pdf (accessed March 18, 2015).

Kolko, J. 2012. *Wicked problems: Problems worth solving: A handbook and a call to action*. Austin, TX: Austin Center for Design. https://www.wickedproblems.com/1_wicked_problems.php (accessed March 18, 2015).

Kurti, M. K., K. von Lampe, and D. E. Thompkins. 2012. The illegal cigarette market in a socioeconomically deprived inner-city area: The case of the South Bronx. *Tobacco Control* 22(2):138–140.

Lasser, K., J. W. Boyd, S. Woolhandler, D. U. Himmelstein, D. McCormick, and D. H. Bor. 2000. Smoking and mental illness: A population-based prevalence study. *JAMA* 284(20):2606–2610.

Lee, J. G., G. K. Griffin, and C. L. Melvin. 2009. Tobacco use among sexual minorities in the USA, 1987 to May 2007: A systematic review. *Tobacco Control* 18(4):275–282.

Lee, J. G., C. B. Agnew-Brune, J. A. Clapp, and J. R. Blosnich. 2013. Out smoking on the big screen: Tobacco use in LGBT movies, 2000–2011. *Tobacco Control* 23(e2):e156–e158.

Leischow, S. J., D. A. Luke, N. Mueller, J. K. Harris, P. Ponder, S. Marcus, and P. I. Clark. 2010. Mapping U.S. Government tobacco control leadership: Networked for success? *Nicotine & Tobacco Research* 12(9):888–894.

Luke, D. 2013. *Systems science and tobacco control*. Presentation at 13th annual American Academy of Health Behavior Scientific Meeting, March 17-20, 2013, Santa Fe, NM. http://aahb.wildapricot.org/Resources/Documents/Luke%20AAHB%202013%20Tobacco%20Control%20_%20Systems.pdf (accessed December 12, 2013).

Macal, C. M., and M. J. North. 2010. Tutorial on agent-based modelling and simulation. *Journal of Simulation* 4(3):151–162.

Malarcher, A., S. R. Dube, L. Shaw, S. Babb, and R. Kaufmann. 2011. Quitting smoking among adults —United States, 2001–2010. *Morbidity and Mortality Weekly Report* 60(44):1513–1519.

Mayhew, K. P., B. R. Flay, and J. A. Mott. 2000. Stages in the development of adolescent smoking. *Drug and Alcohol Dependence* 59(Suppl 1):S61–S81.

McLeroy, K. R., D. Bibeau, A. Steckler, and K. Glanz. 1988. An ecological perspective on health promotion programs. *Health Education Quarterly* 15(4):351–377.

Mermelstein, R., S. Cohen, E. Lichtenstein, J. S. Baer, and T. Kamarck. 1986. Social support and smoking cessation and maintenance. *Journal of Consulting and Clinical Psychology* 54(4):447–453.

Miner, C. 2012. *Information technology, tools, tobacco products, and public health*. http://blogs.fda.gov/fdavoice/index.php/2012/08/information-technology-tools-tobacco-products-and-public-health (accessed February 15, 2015).

Morley, C. P., K. M. Cummings, A. Hyland, G. A. Giovino, and J. K. Horan. 2002. Tobacco institute lobbying at the state and local levels of government in the 1990s. *Tobacco Control* 11(Suppl 1):i102–i109.

NAAG (National Association of Attorneys General). 1998. *Master settlement agreement between 46 state attorneys general and participating tobacco manufacturers*. http://www.naag.org/assets/redesign/files/msa-tobacco/MSA.pdf (accessed April 2, 2015).

NACS (The Association for Convenience & Fuel Retailing). 2014. *House subcommittee examines implementation of Tobacco Control Act*. http://www.nacsonline.com/news/daily/pages/nd0409142.aspx#.VGEsk_nF_OH (accessed November 10, 2014).

NCAI (National Congress of American Indians). 2015. Examples of Tribal Tax Codes, Tax Agreements & Other Resources. http://www.ncai.org/initiatives/partnerships-initiatives/ncai-tax-initiative/examples-of-tribal-tax-codes-tax-agreements-other-resources (accessed March 2, 2015).

NCI (National Cancer Institute). 2007. *Greater than the sum: Systems thinking in tobacco control*. Tobacco Control Monograph No. 18. Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health, National Cancer Institute.

NRC (National Research Council). 2007. *Models in environmental regulatory decision making*. Washington, DC: The National Academies Press.

———. 2014. *Advancing land change modeling: Opportunities and research requirements*. Washington, DC: The National Academies Press.

———. 2015. *Understanding the U.S. illicit tobacco market: Characteristics, policy context, and lessons from international experiences*. Washington, DC: The National Academies Press.

Ockene, J. K., K. M. Emmons, R. J. Mermelstein, K. A. Perkins, D. S. Bonollo, C. C. Voorhees, and J. F. Hollis. 2000. Relapse and maintenance issues for smoking cessation. *Health Psychology* 19(1 Suppl):17–31.

OFR (Office of the Federal Register). 2014. *A guide to the rulemaking process*. https://www.federalregister.gov/uploads/2011/01/the_rulemaking_process.pdf (accessed November 21, 2014).

Plsek, P. E., and T. Greenhalgh. 2001. The challenge of complexity in health care. *BMJ* 323(7313):625–628.

Pokhrel, K., and J. Caine. 2012. *Tobacco control in LGBT communities*. Washington, DC: Legacy.

Prochaska, J. O., C. A. Redding, and K. E. Evers. 2008. The transtheoretical model and stages of change. In *Health behavior and health education: Theory, research, and practice*, 4th ed, edited by K. Glanz, B. K. Rimer, and K. Viswanath. Hoboken, NJ: John Wiley & Sons. Pp. 97–122.

Rittel, H. W., and M. M. Webber. 1973. Dilemmas in a general theory of planning. *Policy Sciences* 4(2):155–169.

Sallis, J. F., N. Owen, and E. B. Fisher. 2008. Ecological models of health behavior. In *Health behavior and health education: Theory, research, and practice*. 4th ed, edited by K. Glanz, B. K. Rimer, and K. Viswanath. Hoboken, NJ: John Wiley & Sons. Pp. 465–486.

Samuel, K. A., K. M. Ribisl, and R. S. Williams. 2012. Internet cigarette sales and Native American sovereignty: Political and public health contexts. *Journal of Public Health Policy* 33(2):173–187.

SFATA (Smoke Free Alternatives Trade Association). 2015. *About SFATA*. http://sfata.org/about-sfata (accessed January 22, 2015).

Smith, E. A., N. Offen, and R. E. Malone. 2006. Pictures worth a thousand words: Non-commercial tobacco content in the lesbian, gay, and bisexual press. *Journal of Health Communication* 11(7):635–649.

Smith, T. T., A. F. Sved, D. K. Hatsukami, and E. C. Donny. 2014. Nicotine reduction as an increase in the unit price of cigarettes: A behavioral economics approach. *Preventive Medicine* 68:23–28.

SNL (Sandia National Laboratories). 2014. Presentation to the Committee on Assessment of Agent-Based Models to Inform Tobacco Product Regulation. Presented at first meeting of the Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation, February 26, 2014, Washington, DC.

Stevens, P., L. M. Carlson, and J. M. Hinman. 2004. An analysis of tobacco industry marketing to lesbian, gay, bisexual, and transgender (LGBT) populations: Strategies for mainstream tobacco control and prevention. *Health Promotion Practice* 5(3 Suppl):129S–134S.

TCLC (Tobacco Legal Consortium). 2009. *Federal regulation of tobacco: Impact on state and local authority*. http://publichealthlawcenter.org/sites/default/files/resources/tclc-fda-impact.pdf (accessed November 13, 2014).

———. 2014a. *A deeming regulation: What is possible under the law*. http://publichealthlawcenter.org/sites/default/files/resources/tclc-fs-deeming-reg-what-is-possible-2014.pdf (accessed November 13, 2015).

———. 2014b. *The FDA takes small steps to expand tobacco regulation*. http://publichealthlawcenter.org/sites/default/files/resources/tclc-fs-fda-takesmallstepexpandtobaccoregulation-2014.pdf (accessed November 13, 2014).

———. 2014c. *Litigation update: Recent cases that implicate the Family Smoking Prevention and Tobacco Control Act*. http://publichealthlawcenter.org/sites/default/files/resources/tclc-fs-litigationupdate-TobaccoControlAct-Sept.2014_0.pdf (accessed November 21, 2014).

TPSAC (Tobacco Products Scientific Advisory Committee). 2011. *Menthol cigarettes and public health: Review of the scientific evidence and recommendations*. Center for Tobacco Products, U.S. Food and Drug Administration. http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/TobaccoProductsScientificAdvisoryCommittee/UCM269697.pdf (accessed April 2, 2015).

Traynor, M. P., M. E. Begay, and S. A. Glantz. 1993. New tobacco industry strategy to prevent local tobacco control. *JAMA* 270(4):479–486.

Tsoukalas, T., and S. A. Glantz. 2003. The Duluth clean indoor air ordinance: Problems and success in fighting the tobacco industry at the local level in the 21st century. *American Journal of Public Health* 93(8):1214–1221.

TUDGP (Tobacco Use and Dependence Guideline Panel). 2008. *Treating tobacco use and dependence: 2008 update*. Rockville, MD: U.S. Department of Health and Human Services.

Washington, H. A. 2002. Burning love: Big tobacco takes aim at LGBT youths. *American Journal of Public Health* 92(7):1086–1095.

Westmaas, J. L., J. Bontemps-Jones, and J. E. Bauer. 2010. Social support in smoking cessation: Reconciling theory and evidence. *Nicotine & Tobacco Research* 12(7):695–707.

Young, D., R. Borland, and K. Coghill. 2012. Changing the tobacco use management system: Blending systems thinking with actor–network theory. *Review of Policy Research* 29(2):251–279.

Ziedonis, D., B. Hitsman, J. C. Beckham, M. Zvolensky, L. E. Adler, J. Audrain-McGovern, N. Breslau, R. A. Brown, T. P. George, J. Williams, P. S. Calhoun, and W. T. Riley. 2008. Tobacco use and cessation in psychiatric disorders: National Institute of Mental Health report. *Nicotine & Tobacco Research* 10(12):1691–1715.

3

Building Effective Models to Guide Policy Decision Making

Although policy makers have long looked to behavioral models to guide their decision making, there is no accepted set of recommendations or best practices for how to manage this process. In accordance with its statement of task, the committee reviewed the uses of agent-based modeling (ABM) in policy decision making and how this method fits into a broader methodological toolkit. The goal of this chapter is to provide guidance on (1) understanding the conditions under which models—specifically individual-level models—are appropriate and useful in aiding policy decisions; (2) elucidating the empirical and theoretical challenges of specifying model inputs and interpreting model outputs appropriately; and (3) providing guidance for navigating key modeling decisions, including determining the appropriate levels of verisimilitude and aggregation, dealing with issues of model specification and evaluation, and quantifying uncertainty. Fortunately for tobacco control policy modelers, many regulatory authorities and academic fields are struggling with related problems in terms of model specification and inference. Their efforts offer a wealth of examples and experiences to draw from.

The organization of this chapter is as follows. The motivation for models in policy decision making is described. The committee articulates specific mechanisms through which human behavior may depend on the behavior of others as well as on features of the local environment. Then the major challenges to getting empirical evidence to adjudicate among these alternative mechanisms are reviewed. Next, a number of key distinctions in modeling are introduced, including micro- versus macro-level models, analytical versus computational models, and models that incorporate varying levels of detail in representing a given process. The appropriateness of each type of model under different levels of uncertainty and data availability is discussed. The committee suggests methodological strategies for specifying individuals' behaviors within micro-level models and for assessing how uncertainty in model inputs translates into uncertainty in model outputs.

THE CHALLENGE OF ANTICIPATING AND UNDERSTANDING POLICY EFFECTS

Policies can backfire when they fail to account for how people change their behavior in response to an intervention. This is known as "policy resistance" in the public health literature (Sterman, 2006) and "blowback" in covert operations. It goes back to old social science literature on the "law of unintended consequences" (Merton, 1936; Smith, 1759). The basic issue is that individuals' behavior often depends on the behavior of other people or features of the social environment, or both. Any policy that aims to change behavior or outcomes can result in a chain reaction of events that can potentially undermine the efficacy of that policy.

This problem arises in many substantive areas. To take an example from tax policy, if workers allocate their time to maximize both earnings and leisure, an overly stringent income tax may lead them to cut back on hours worked, which may in turn reduce total government revenue from taxes (Saez et al., 2012). Within the domain of transportation, antilock brakes can cause people to drive more aggressively, thus partially offsetting their safety benefits (Wilde, 2001). Closer to home for readers of this report, there is evidence that low-tar and low-nicotine cigarettes may actually increase the intake of carcinogens, as people smoke more frequently and hold the smoke in their lungs for longer (HHS, 2010; NCI, 2001).

Although, as the above examples show, a policy may generate negative feedbacks, positive feedbacks may also occur, enhancing the effectiveness of the policy. In the classroom, the provision of tutoring or other special help to some students may indirectly aid the learning of other students as members of the class interact with one another. Persuading one person to stop smoking may influence friends and family to stop smoking as well. Such positive feedbacks are sometimes called *social multipliers* (Manski, 1993).

Whether feedbacks are negative or positive, a central challenge for policy makers is to anticipate how organizations, corporations, and individuals will react to changes in incentive structures and features of the environment. Anticipating this response can be difficult for a number of reasons. One challenge is that knowledge of human behavior is limited and that it is difficult to infer from past behavior how people will respond to novel situations. A related problem is that people's behavior is both influenced by and also influences the behavior of others, through direct interactions (e.g., social influence and peer effects) as well as features of the social environment. This makes it difficult to assess the global effect of a policy or to anticipate its efficacy at different scales of implementation.

For example, a housing policy that encourages a small number of individuals with low income to move to higher-income neighborhoods may appear to successfully accomplish its intended goal of economic integration. However, if that policy were to be expanded to a larger population, the higher-income residents of those neighborhoods might move out (presumably, because the neighborhood has declined), which in the end would leave these lower-income households no better off than before. Conversely, an antismoking policy targeted at a small group of persons may have little positive effect, but one targeted at a larger group may generate a change in social norms that induces persons not within the target group to stop smoking as well. To be maximally effective, policy makers must be able to assess their proposed interventions' total effect, including how affected individuals, organizations, or institutions might adapt to a new environment or change their behavior in reaction to what others are doing.

Anticipating the Effects of Policies

Historically, the "gold standard" for evaluating the effects of a public health intervention has been an analysis of treatment response using data from randomized controlled trials (RCTs). This approach overcomes the fundamental problem of causal inference: For any given treatment unit, the counterfactual outcome is never observed—that is, what would have happened if that unit had or had not received the treatment. By removing the possibility of selection bias, RCTs provide a more rigorous test of treatment effects than do observational studies.

Information gleaned from RCTs alone is often insufficient for guiding policy decision making. Perhaps the most obvious issue is that it may not be feasible or appropriate to carry out the desired RCTs. This is partly due to practical limitations: It is impossible to design RCTs to test all possible policies. There may also be legal or ethical restrictions that make RCTs inappropriate. In some cases quasi-experimental methods (e.g., instrumental variables) or modeling strategies (e.g., propensity score matching) can be used in an attempt to mimic experimental research design, but these approaches may require one to make implausible assumptions in order to produce inferences.

In addition, RCTs are ill-suited for evaluating policy effects when the behavior of different individuals is interdependent. Indeed, the standard analysis of RCTs makes the assumption that one person's treatment outcome is independent of who else received the treatment. When the efficacy of one person's treatment depends on whether others received the treatment, the methodology falls apart. For example, RCTs have limited ability to inform society about the effectiveness of vaccination policies for a population susceptible to infectious disease. An RCT with a small treatment group might provide information about the payoffs to vaccination when a small number of people are vaccinated, but credibly extrapolating from this to a larger treatment population may prove to be impossible, for two reasons. First, any individual's decision whether to get vaccinated may depend on how many others are getting vaccinated. Second, the danger of catching a disease varies with overall rates of vaccination. An RCT examining the effectiveness of a tobacco use cessation treatment program would have similar problems. The treatment of one individual could have beneficial effects on others—for example, on the individual's spouse or peers, who may quit in reaction to the treated individual successfully quitting.

Finally, traditional analyses of RCTs tell us only what does or does not work; they provide no information on the reasons why an intervention worked or not. Thus, the information gleaned from RCTs and quasi-experimental methods may lack external validity. This makes it difficult to extrapolate the effects of interventions implemented in one context to a different context or to infer the expected effects of novel interventions from knowledge about the effects of prior interventions (Cartwright, 2007; Heckman, 2008; IOM, 2012; Manski, 2013). On the other hand, experimental or quasi-experimental estimates can be used to guide theory and shed light on underlying structural relationships. In a complex world, moving toward a more structural approach—and away from a "black box" analysis of experiments (that is, not having the ability to understand the inner workings of the processes under study)—will help researchers do more than estimate an intervention's causal effect. How a policy is expected to work within the larger social context requires system-level knowledge and a sense of the behavioral mechanisms through which it operates (Sampson et al., 2013).

Structural Models

Structural models use a set of equations or rules—expressed analytically or computationally in programming code—to describe different possible worlds. The specification of the model is dictated by theory, prior knowledge, and other inputs that determine which features of a given process to highlight and which to leave out. These assumptions, combined with data, produce a set of inferences about what will happen under a given set of conditions. This modeling approach includes (but is not limited to) macro-level simulation models, such as system dynamics models, and micro-level simulation models, such as ABMs. The appropriateness of a given modeling strategy depends on the theory brought to bear and on the available empirical evidence.

Structural models typically attempt to capture behavioral relationships or parameters that hold true across a range of social conditions or take those conditions as inputs that affect behavior. This requires a reasonably deep understanding of the incentives that drive behavior. For example, one might observe an association between neighborhood poverty and rates of teenage pregnancy. An example of a superficial model of this process would be one that translates this aggregate correlation into individuals' decisions: As poverty increases, the likelihood of a young woman having a child goes up by some amount. However, this model ignores the underlying motivations for these women's decisions, how other people may influence those decisions (e.g., parents and partners), and how decisions are predicated on these women's beliefs about the benefits and costs of having a child, which may depend on what opportunities are available to them. Without taking into account these underlying motivations for behavior, the hypothetical model is extremely brittle in its ability to make inferences about how women would behave under alternative scenarios. The general point is that the more fundamental—or "deeper"—the relationships captured in a model, the more effective the model is at exploring the implications of a wider set of policy scenarios (Blume, 2015; Heckman, 2008). Thus, an additional criterion for model usefulness is that it has parameters that are sufficiently fundamental to cover all the policies under consideration (Marschak, 1974).

One challenge in using structural models effectively is specifying them in a way that is empirically defensible and that allows for a clear and rigorous quantification of the assumptions embedded in the model (NRC, 2014). To be useful, a model must be able to quantify how uncertainty in the model's inputs translates into uncertainty in the model's outputs (Manski, 2013). Structural models vary widely in how complicated they are. Models that include more parameters and greater verisimilitude do not necessarily make more assumptions than simpler models. This is because in many cases researchers can conceive of models with more parameters than there are available empirical inputs. Regardless of model complexity, models are more credible if parameters are backed up by hard evidence or, at minimum, a well-developed theory. Whatever the level of detail used to represent a process, models that guide policy must be both credible and sufficiently explicit in their assumptions about the process under investigation.

The degree of model verisimilitude may reflect different goals. Some models are designed to run virtual experiments to determine the outcomes that could be expected from implementing different policies. Other models have a simpler goal: to identify the potential pitfalls or unanticipated consequences of a given policy or to get some sense of what RCT design would be needed to accurately assess policy effects. In both cases, if the models are to produce valid inferences, they must be able to capture accurately the distribution of outcomes that might be expected under a given set of conditions and to suggest which outcomes are more or less likely. For some policy makers and model developers, one attraction of ABM is that it allows for almost unlimited detail in representing the process under investigation. More complicated models do not necessarily generate more accurate predictions, especially if data, theory, and other model inputs are insufficient to identify the foundational parameters of the model (Sanstad, 2015).

MECHANISMS THAT GENERATE FEEDBACK BETWEEN BEHAVIOR AND SOCIAL ENVIRONMENTS

There are several different ways that people's actions can be influenced by their environment, which includes both what others are doing ("social interactions") and the institutional, political, and organizational factors that shape people's incentives, such as the regulatory environment. Policies may be more effective when they can directly target the specific mechanism that gives rise to the process under investigation, and thus policy makers need to evaluate an ABM's ability to explicate the behavioral mechanism under investigation. This section reviews the theoretical and empirical literature on mechanisms governing contingent behavior and suggests some ways in which these insights might be fruitfully applied in the domain of tobacco regulation. Note that here the focus is on mechanisms that occur "above the skin" (for example, environmental or societal factors). For a review of structural models that attempt to capture interactions "below the skin" (for example, genetic, metabolic, and neurobiological factors), see Hammond (2015).

Social Interactions

People's behavior is often shaped by what others are doing. This type of phenomenon is often referred to as social interactions, social influence, or spillover effects. Manski (2000) distinguishes between three types of social interactions. First, there are *constraint interactions*, which cause an action to become less desirable and available as more and more individuals engage in it. One example of this is freeway congestion: Freeway driving is most attractive when there are few people on the road and increasingly less desirable as more and more people use the highway. Second, *expectations interactions* occur when people draw inferences about expected outcomes of a given action or about difficult-to-observe attributes of a situation or person based on prior experience or an outside body of knowledge. One case of this is statistical discrimination. For example, an employer may have certain expectations about young workers—that they are more likely to quit their jobs in order to go back to school—and this influences the employer's enthusiasm for hiring from this population. Similarly, teenagers may observe the effects of smoking on older relatives, which shapes their beliefs about the effects of tobacco on health. Finally, *preference interactions* occur when a person's ordering of attractiveness concerning some set of alternatives depends on the choices of others. For example, in the case of "white flight," each time white persons leave a neighborhood because they cannot tolerate the presence of minorities, they leave the neighborhood a bit less white behind them, thereby inducing other whites to exit as well (Schelling, 1978).

These are theoretically distinct processes, each of which suggests different policy interventions, but in practice it is difficult to empirically distinguish them. Moreover, although the cases outlined above represent different instances of *endogenous effects*, people who share the same social context may display similar behavior even in the absence of these social interactions. For example, similar behavior may arise from *contextual effects*, which refer to the way in which people's behavior is shaped by a shared social environment, such as neighborhood composition or school quality. Also, a group of people may share similar behavior or outcomes due to *correlated effects*, which refer to a situation in which people share the same attributes or opportunities. For example, people within the same birth cohort may have similar career trajectories, on average, in large part because they face the same job market conditions. A challenge for researchers who believe they have identified some sort of endogenous behavior is to identify the effects of the social influence apart from shared opportunity structure (Manski, 2007).

Imagine a case in which some correlation is observed between peer group membership and whether and how much a teenager smokes cigarettes. There are four ways in which this result might come about. First, students in the same peer group might influence one another's smoking behavior through the availability of cigarettes or through peer pressure, or both. Second, the students in the same peer group may share similar individual attributes or risk factors (e.g., gender, family resources, parents' education) that affect smoking. Third, the students may affect one another's smoking behavior through behavior other than their own smoking behavior. For example, if students in the peer group are more likely to cut class, and if cutting class leads to higher rates of smoking, a correlation between peer group membership and smoking could be observed. Finally, students may inform one another about the existence and properties of different forms of tobacco (e.g., cigarettes, e-cigarettes, hookah, chewing tobacco). These different pathways would suggest different policy interventions. In the second case, where behavior is a function only of individuals' attributes, there are no social interactions.

One difficulty in identifying social interactions stems from the fact that the average behavior within a group is itself a function of the behavior of group members. Thus, observing a correlation between peer group membership and smoking behavior does not tell us whether peer groups influence the behavior of their members or the groups' behavior is simply aggregating over the behavior of group members. This is known as the "reflection problem" (Manski, 1993, 2000, 2007). If the data contain sufficient variation between and within groups, it may be possible to determine whether individual attributes alone can explain variation in behavior across groups. Even if researchers have reason to believe that group affiliation shapes behavior, they still must distinguish among the different types of social interactions described in the previous paragraph. In many empirical cases, it is difficult or impossible to identify the groups that are actually influencing behavior. The reflection problem is even more difficult to resolve when group affiliation is unknown. See Manski (1993, 2000) and Blume et al. (2010) for more detailed discussion of this issue.

An example of the challenge of identifying the presence of social interactions—let alone determining the nature of those interactions, if they exist—can be found in the debates over the effects of peer influence on smoking. Christakis and Fowler's (2008) study examines whether knowing people who quit smoking makes it more likely that a given individual will also quit smoking. Their results suggested that friends, coworkers, siblings, and spouses had dramatic effects on adults' smoking behavior. For example, they found that a person is two-thirds more likely to quit smoking if his or her spouse also quits. A coworker, sibling, or friend quitting had a smaller, but nontrivial effect—ranging from one-quarter more likely to quit smoking in the case of a sibling to over one-third in the case of a friend. Identifying true "social contagion" effects requires separating out the effects of *homophily* (people's tendency to select others who resemble them on observed or unobserved attributes) and *shared social environment* from the effects of *social influence* (Aral et al., 2009; Shalizi and Thomas, 2011). Indeed, later studies that used more rigorous strategies for controlling for unobserved features of the environment and selection artifacts have found that peers have far more modest effects on smoking behavior (e.g., Fletcher, 2010; Fletcher and Ross, 2012). One lesson here is that policies aimed at encouraging or discouraging the spread of behaviors in networks must be backed by rigorous empirical studies that convincingly separate homophily effects from effects due to social contagion (Aral et al., 2009; Shalizi and Thomas, 2011).

More generally, to be useful for informing regulatory policy, modeling efforts must capture meaningful aspects of the social process under investigation. If the goal is to understand how interdependent human behavior will shape the outcomes experienced under a given policy, a serious empirical effort is required to determine the underlying mechanism at work. It is not enough to hypothesize different mechanisms and use a model to determine whether those mechanisms lead to different outcomes. The model may be misspecified to the point where a "sensitivity analysis"[1] provides no information at all on the true sensitivity of model outputs to inputs (Sanstad, 2015).

Institutional Factors

This chapter has focused on how individuals' behavior influences the behavior of others, but there are also higher-level structures (e.g., tax policy or safety regulations) that shape people's choices. There is evidence across a number of policy domains that if the incentives or risks associated with a given behavior are changed, people will likely behave differently. For example, some evidence suggests that the development of highly effective HIV treatments has been associated with an increase in unprotected sex among people living with HIV in the United States (Katz et al., 2002; Lightfoot et al., 2005).

The behavior of organizations and other coalitions is also influenced by behavioral incentives. Failure to account for those incentives may lead to unexpected and undesirable results. For example, an increase in U.S. corporate taxes may result in some firms decamping for countries with lower tax rates, thereby reducing total U.S. tax revenue (Devereux and Maffini, 2007). Similarly, tobacco companies make strategic decisions that are influenced by the current regulatory environment and will work to counteract the efficacy of policies aimed at reducing smoking. Congress or states may change taxes on cigarettes, for instance, but the tobacco companies may respond by offering coupons or bulk discounts (Arno et al., 1996; Henriksen, 2012; Loomis et al., 2006). A model that aims to predict how people's or firms' behavior might change under a different incentive structure must therefore understand the reasons for their behavior. Later in this chapter, strategies for specifying models of behavior that try to account for these motivational factors are discussed.

> *Conclusion 3-1: The committee concludes that a deep understanding of human behavior, decision making, and incentive structures is important for agent-based models and other models that are used to understand how interdependent behaviors shape the outcomes of a given policy. Regardless of the model type, if the behavior is not plausible, the model is not likely to be informative.*

---

[1] Sensitivity analysis is "an exploration, often by numerical (rather than analytical) means, of how model outputs (particularly QOIs [quantities of interest]) are affected by changes in the inputs (parameter values, assumptions, etc.)" (NRC, 2012, p. 117).

This conclusion is especially relevant when a model is intended to explore which outcomes might occur (or how people would behave) under alternative scenarios.

There is some debate about how plausible a model's representation of individual behavior must be in order for that model to be informative. Friedman (1953) argued that models do not need to represent the underlying process accurately to be useful—the model only has to predict well. A practical problem with that line of thinking is that it presumes the existence of evidence that the implausible model actually predicts well—but how would this be known ex ante? It can only be discovered ex post. Another problem is that models that do not accurately represent the underlying process under investigation at some degree of fidelity can be brittle, losing their predictive power when conditions or incentives change. This is in contrast to structural models that capture key features of a process (Marschak, 1974).

Sometimes models that inaccurately represent individual behavior yield qualitatively accurate aggregate predictions. For example, Schelling's model (1978) of how individual decisions generate aggregate patterns of segregation includes several implausible assumptions about behaviors (for example, that people make decisions about where to live based on whether their own racial group is the local minority or majority and that there is no cost to moving). The Schelling model did provide the important theoretical and policy-relevant qualitative insight that segregation can emerge even though people have preferences for racial integration. However, it does not give a credible quantitative prediction, and so it does not provide a suitable basis for predicting when segregation will emerge in specified real-world settings. Moreover, it would be a mistake to use the Schelling model to predict how households might respond to pricing incentives, counseling, or other interventions aimed at promoting neighborhood diversity.

Conclusion 3-1 is most relevant when models will be used to inform policy decisions. Models that do not include an understanding of human behavior, decision making, and incentive structures can be informative for some purposes. However, it is the committee's view that models developed for the purpose of anticipating the effects of policy decisions need to have some anchoring in real-world behavior.

> **Recommendation 3-1: When developing an agent-based model (or similar modeling approach), the Center for Tobacco Products should consult with subject-matter experts to identify the plausible behaviors and focal processes at work from the beginning of the model development process.**

Chapter 4 also discusses the need for input from subject-matter experts throughout the lifespan of a model. An essential feature of ABMs is the representation of agents in the real world, and because agents often have distinct characteristics and behaviors, non-experts can inadvertently misrepresent agent behavior. This makes collaboration with subject-matter experts essential at all stages of model development.

OVERVIEW OF TYPES OF STRUCTURAL MODELS FOR INFORMING POLICY DECISIONS

Thus far, the committee has discussed the role that models can play in guiding policy decisions, and it has reviewed some of the behavioral mechanisms that can lead to feedback between individuals' behavior and the social and regulatory environments. This section provides a high-level overview of the types of models that are used to capture this type of feedback behavior. A number of key features of such models are reviewed, including whether they have an analytical versus computational solution and whether they capture phenomena at the individual or group level. The chapter also addresses the confusion regarding the distinction between microsimulation and ABM, which the committee believes limits researchers' and policy makers' ability to incorporate lessons learned and best practices from the array of studies in different disciplines. This section ends with a discussion on how to define the appropriate level of model specificity and how to anticipate and understand the different forms of equilibrium outputs generated from models.

Analytical and Computational Models

Recall that a structural model is a set of equations or rules for how the individuals or other units in a specified population interact and are influenced by their environment, and it can be implemented both analytically and computationally. The National Research Council defines a computational model as "computer code that (approximately) solves the equations of the mathematical model" (2012, p. 110), whereas analytical models can be solved mathematically in closed form, that is, the solution to the equations used to describe changes in a system can be expressed as a mathematical analytic function (NRC, 2007).

If the behavioral rules or the population structure is sufficiently simple, it may be possible to determine analytically (i.e., mathematically) how the state of the population changes over time and whether the population gravitates toward a steady state (equilibrium). However, if the rules are sufficiently complex or the population is too heterogeneous, it may be impossible to determine the dynamics of the system or to derive the steady state. In some cases, there is no steady state solution (e.g., Salop and Stiglitz, 1982). In this case, the analyst must simulate the process iteratively by applying the rules and updating the population composition to arrive at a final solution or observe the dynamics. Note, however, that some models only determine equilibrium outcomes without specifying the process by which the social system attains equilibrium. For these models, intermediate solutions have no substantive meaning, and thus there are no dynamics.

Even seemingly minor relaxations of assumptions may make it impossible to solve a model analytically. For example, a simple model of disease spread that assumes that interactions among individuals are equally likely ("random mixing") can be solved mathematically. As soon as one relaxes this assumption and allows for different individuals to have varying rates of exposure, the heterogeneity in agents' disease risk makes the resulting model analytically intractable (Blume, 2015; Osgood, 2007). Thus, the results must be simulated.

In general, analytical models have relatively simple specifications of behavioral rules that assume tractable forms of interactions (Grazzini et al., 2013). Models with analytical solutions are often more restrictive than simulation models, as they have fewer parameters and simpler interaction structures. This is not necessarily bad. Analytical models have several advantages over computational models. First, because the equilibrium solutions are derived mathematically, it is possible to identify the whole space of solutions. This is particularly important when there exist *multiple equilibria* for the same set of model inputs, as is often the case with social interactions models. (This point is discussed in greater detail in the next section.) For policy makers it is of great interest to identify the potential for multiple equilibria, as this suggests that, insofar as the model has captured fundamental features of the process, the same starting conditions might, depending on how a process unfolds, end up in very different final states. A possible policy goal may move from a "low level" equilibrium to a "higher level" equilibrium (Moffitt, 2001). Feedback suggests the possibility that, with sufficient understanding of incentive structures, one might "harness" the interactions so that there are bigger payoffs relative to costs. In addition, analytical models can also reveal the path to equilibrium, which may be more important than the equilibrium itself. For example, if the goal of the model is to anticipate outcomes over a finite time horizon, knowing the equilibrium outcome is of little use if it applies to a world that exists decades or centuries into the future.

Second, analytical models allow the analyst to determine the stability of model equilibrium—in other words, how likely it is that the model will return to a given state if it is slightly perturbed away from that state. Whether or not an equilibrium solution is stable may have important policy ramifications. Regardless of the attractiveness of a particular outcome from a social welfare standpoint, if that outcome is highly unstable, it may be impossible to maintain it in real-world situations. Moreover, understanding the "gravitational pull" of different equilibrium solutions provides insight into how the process under investigation may translate from one stable state to the next. Thus, evaluations of policy outcomes must take into account not only the attractiveness of a given result but also the likelihood of maintaining it over time. Analytical models allow researchers and policy makers to take both factors—which equilibria are possible and which are most likely to be sustained—into account. This is much harder to do with simulation models, which may not identify highly unstable equilibrium solutions.

The potential downside of analytical models is that they may only be able to represent a small number of features of a real-world setting and may make simplifying assumptions that reduce the empirical plausibility of the process represented. However, it is a mistake to assume that simply adding more features to a model will provide a better representation of the process under consideration. This is especially true if there is a great deal of uncertainty in how those features should be specified. Researchers need to be clear about what they are giving up for the benefits of added verisimilitude. Users of simple models may be able to understand model behavior very thoroughly, whereas users of more complicated models may lose their grasp on how a given set of results came to be. Therefore, one would only move to a complicated model if the simpler form is understood and there is a reasonably clear and accurate empirical representation of the more complex process.

It is often useful to start with an analytical model and then expand on it slowly, making effort to tie results from computational solutions to their simpler foundations. Examples of this approach include Brown et al.'s (2004) analysis of green belts and Epstein et al.'s (2008) analysis of the coupled spread of disease and fear about the disease. In these cases the researchers took pains to try to understand completely the simple dynamics involved before turning to more complicated and realistic simulations. Furthermore, modelers may use both analytical and computational methods to describe the relationships between individual choice making and aggregate outcomes. For example, a model may be solved analytically to determine the optimal behavior of each individual agent conditional on the behavior of other agents, but then solved using simulation to determine the equilibrium outcome among many agents.

Equilibrium

As mentioned in the previous section, the focus of a model may be to predict the steady states (equilibria) of a system or to predict its dynamics out of equilibrium. When is it useful to focus on out-of-equilibrium versus equilibrium predictions? It depends on how stable an equilibrium is and how long the system being modeled takes to reach the equilibrium. Consider the metaphor of a rocking chair. If the rocking chair is perturbed, it might start out rocking quickly but then settle into a steady state or equilibrium. Because this happens fairly quickly, it may be valuable to develop a model that focuses on equilibrium conditions.

However, it may take a very long time—perhaps decades—for a real social system to reach equilibrium. In this case a model that focuses on equilibrium would be useless, and it becomes important to be able to credibly predict the dynamics that would follow an intervention. Imagine, for instance, that the U.S. Food and Drug Administration initiates an information campaign. People learn something new about tobacco, and their resulting behavior changes in turn affect other people. In this case, the dynamics of social learning would unfold gradually. If the process took only a few months to reach equilibrium, it might suffice to analyze only the equilibrium conditions. On the other hand, if it would take 50 years for the dynamics to play out, then an equilibrium model would be less useful. Thus it is important to consider the speed at which equilibrium is reached, as this has policy implications. For tobacco control policy, not much is known about the time scales over which equilibria may be reached. An example of when it might make sense to examine only equilibrium conditions is a model of the effect of price on smoking prevalence, which falls rapidly following a price increase. Several econometric models have been developed to estimate the final effect of such a price hike (Chaloupka and Pacula, 1999; Chaloupka and Warner, 2000). These models do not try to represent how smoking prevalence changes over time; they focus only on the final value.

Of course, exploring the time to equilibrium requires that the model be initialized in some starting condition that is anchored empirically. Moreover, the model needs to have a meaningful time scale so that "model time" may be mapped onto "real time."

Individual-Level and Aggregated Models

Another decision that must be made in the modeling process is whether the basic units of the model will be at the level of individuals or aggregated groups. The same process may be represented at different levels of aggregation. For example, one might specify a model of teenage smoking behavior that assumes that school attributes influence girls' and boys' smoking decisions differently but that all girls and all boys have the same response to the environment. In this case, one could specify a model that represents the process for girls and boys separately, but not for individual children. Or, by contrast, the analyst could allow for each child to have a unique set of inputs in the decision process and model the process at the individual level.

Both approaches have strengths and weaknesses. Aggregate models are often easier to build and interpret, but they can only handle a limited amount of population heterogeneity. If population heterogeneity is a key feature of the process under consideration, or if the model incorporates individual-specific trajectories or experiences (e.g., work histories), the analyst will likely need to specify the model at the individual level in order to allow each person to have a unique profile.

From an implementation standpoint, aggregate models have certain advantages over individual-level models. They are more straightforward to construct and understand, and they often take less time and computational power to run. Finding empirical data to anchor them may also be easier. In addition, if the analyst wants to simulate the dynamics of a very large population (e.g., the population of the United States), individual-level models can easily become unwieldy. Researchers have to weigh the trade-offs.

Both aggregate and individual models can incorporate feedback effects across levels of analysis. However, it is difficult to incorporate social interaction effects into aggregate models if these effects occur at the local level. (If there are global interactions, for example, where all individuals respond to the total number of people working in the population, individual-specific response functions are not required.) A key challenge in implementing individual-level models is finding the empirical knowledge or data necessary to make them credible. The data demands for an individual model are higher than those for an aggregate level model, especially in terms of the plausible specification of individuals' behavior (see Chapter 6 for a discussion on data needs).

Microsimulation and Agent-Based Models

Within the domain of individual models, some scholars distinguish semantically between two types of models: *microsimulation* and *agent-based models*. Both involve the same basic procedure: Artificial agents are assigned a behavior, and simulation is used to assess the aggregate implications of that behavior. Both modeling approaches are operationalized through computer code. Thus, in theory, anything that is specified as an ABM can be specified as a microsimulation, and vice versa. It is important to note this commonality because, if viewed as two distinct approaches, the two research communities are less likely to benefit from each other's work. However, some argue that there are a number of differences between ABM and microsimulation, both in the research questions they consider and in their common practices.

For example, in a review of the literature on ABM, Macy and Willer (2002) claimed that the difference between ABM and microsimulation is that microsimulations assume no interaction among agents. And, indeed, many microsimulations do attempt to explore how heterogeneous populations respond to some change in policy or incentives, without allowing for interactions among agents. For instance, the Congressional Budget Office's (2007) microsimulation tax model explores how the U.S. population might respond to a change in tax rates, taking into account the fact that different types of people (for example, married and unmarried, men and women) have differential responses. This model does not specify that agents interact; rather, its goal is to compute the net response, taking into account the fact that people's labor force participation is contingent on their expected income after taxes. However, the committee found many examples of microsimulations in which the environment of the agents is generated from agents' previous decisions. As one example, Mare and Bruch (2003) used a microsimulation to determine the equilibrium segregation outcomes implied by agents in a residential mobility model.

Although from a purely technical standpoint microsimulations and ABM are the same modeling enterprise, the committee did find differences in how these techniques tend to be deployed. For instance, microsimulation models typically keep the agents' environments simple and abstract, as these models are anchored in even simpler analytical models for which the dynamics are well understood. ABMs are sometimes grounded in analytical models, but this is not standard practice. Also, ABMs may incorporate highly detailed environments in which the agents interact, drawing on maps and other geographic information. This is technically possible with microsimulations, but in general microsimulations tend to abstract away from spatial features of the agents' environments. In addition to their emphasis on spatial interactions, ABMs tend to emphasize other features of complex systems, including population heterogeneity, adaptation, and learning (Hammond, 2015). In short, there are not fundamental differences between ABMs and microsimulations, but there are historical differences in how these models have been specified and used by their research communities.

> *Conclusion 3-2: Researchers who use the terms agent-based modeling and microsimulation have different approaches to model specification. However, the committee concludes that from a technical standpoint these are the same enterprise (an individual-level model implemented via computer code). The committee believes that modelers would greatly benefit from best practices and lessons learned from applications that have been performed by the two research communities to address policy questions.*

This report is focused on ABM, and it is the committee's sense that agent-based modelers would benefit from drawing on the large literature on microsimulations, especially in the context of policy decision making. For example, microsimulations have been used in tobacco control in recent years, and CTP and other agent-based modelers could look to those examples (Jeon et al., 2012; van Meijgaard et al., 2009). Thus, in the remainder of the report these two methods will be treated as technically the same approach, albeit with different implementation practices.

High-Dimensional Models and Low-Dimensional Models

Finally, as noted earlier, the model developer must decide on the appropriate level of the model's detail and empirical realism. The appropriate level of model detail depends on the research question, the intended use of the model, and the data that are available to empirically ground the model. It is important to note, however, that no matter what level is chosen, models provide only an imperfect representation of the real world, as computational models in general are not reality mirrors, nor are they intended for this purpose. ABMs can represent anything from low-dimensional, abstract worlds where agents are defined by just one or two attributes and interact in a highly stylized environment based on simple rules, to high-dimensional, highly detailed worlds where agents have many attributes, the environment contains a great deal of information, and agents engage in multiple behaviors (Bruch and Atwell, 2013).

It may be tempting to design an ABM that pulls in all the empirical data and knowledge available in order to create a highly realistic "laboratory" in which to explore policy questions. However, this approach is not usually the most productive, especially at the early stages of modeling, as the available data and knowledge of human behavior are generally not available. While data on demographic, biological, and social characteristics of individuals, families, or other groupings are often collected, data on how those units interact are generally lacking. ABM allows the developers to explore how important various mechanisms are when data are lacking and to assess the potential value of collecting these data; however, this introduces an added layer of uncertainty and raises the possibility of model misspecification. Furthermore, the model can become cumbersome and hard to manage when additional layers of detail are added, and it can be difficult to get clear analytic results (Blume, 2015). The success of a model is not determined by the level of granularity at which it represents a process; rather, its success is based on how successfully it facilitates the understanding of the problem or question under study.

Conclusion 3-3: The committee concludes that low-dimensional and high-dimensional models have complementary virtues and weaknesses. A more complicated model may have greater verisimilitude, but added detail per se does not ensure realism. A low-dimensional model, while abstracting from some features of the real world, may generate forecasts that are easier to understand and interpret.

Recommendation 3-2: The Center for Tobacco Products should develop and employ both low- and high-dimensional models, using both as appropriate to shed light on policy impacts.

SPECIFYING INDIVIDUAL BEHAVIOR IN AGENT-BASED MODELS

This chapter began with a discussion of why policy makers need empirically grounded models to anticipate the effects of their policies. However, those models are only useful insofar as they accurately capture what outcomes would occur under alternative scenarios. A major factor in evaluating the credibility of a micro-level model is whether or not it has captured the core behaviors of individuals or organizations or other units under investigation. This is especially important when the only data available to understand people's response come from a population in which the focal policy has not yet been implemented or has only been implemented on a small scale. Analysts need some way of making empirically defensible claims about how people might change their responses under different conditions. This section discusses different approaches for specifying individual behavior within simulation models. Although a reasonable specification of behavior may not be sufficient for generating a useful model, it is necessary for valid inferences. The point of structural models is to capture fundamental features of the process under investigation. If individual behavior is misspecified, particularly in an individual-level structural model, it is difficult to see the value in the enterprise.

Quantitative Approaches

One approach to specifying individual behavior is to postulate that agent preferences or behaviors are captured by the parameters of a quantitative model. If the behavior under investigation involves discrete changes in agents' attributes—for example, marriage or childbirth—these transitions can be described using coefficients from a discrete-time event history model (Allison, 1982). If the behavior under investigation implies some sort of decision process (for example, the decision to smoke), discrete choice statistical models provide a useful framework for developing an empirically grounded representation of agents' choice behavior (Ben-Akiva and Lerman, 1985; McFadden, 1974). Historically, discrete choice models have been based on a rational actor model of behavior in which individuals have unlimited computational abilities for performing the calculations necessary for evaluating all possible options.

Discrete choice models have become more behaviorally sophisticated in recent years, drawing on largely experimental work in psychology and decision theory that demonstrates that people have limited time for learning about available options, limited working memory, and limited computational capabilities. These choice models allow for "variation in individuals' knowledge of available options; strategies for learning about or evaluating available options; reactions to change in environmental conditions; reactions to past experiences; and susceptibility to social influence" (Bruch and Atwell, 2013, p. 11). For an example of contemporary discrete choice models that incorporate decision makers' cognitive strategies to reduce the demands of evaluating potential options, see Gilbride and Allenby (2004) and Hauser et al. (2010). However, to the best of the committee's knowledge these models have not been applied to problems outside of marketing, so their value for public health applications remains unknown. Regardless of the choice model used, "estimation of relevant coefficients requires information on either revealed preferences (observed choices) or the stated preferences (survey responses to hypothetical choice scenarios) for some population of interest" (Bruch and Atwell, 2013, pp. 11–12). Surveys, observational data, and administrative records are potential sources for this kind of data.

In recent years, a line of work spearheaded by Brock and Durlauf (Blume et al., 2010; Brock and Durlauf, 2001; Durlauf, 2001) has developed discrete choice models that explicitly model social interactions. In other words, the utility or payoff that a person gets from a particular action depends directly on the characteristics or behavior of others. When the characteristics of other reference group members enter the choice function, this reflects contextual effects, as discussed earlier in this chapter. For example, if the availability of female role models influences women's decisions to major in STEM (science, technology, engineering, and mathematics) fields, the number or proportion of available female role models may be incorporated as a background covariate in the model. Alternatively, these variables may capture endogenous effects whereby individuals' choices are contingent on the choices of others. For example, a teenager's decision to engage in some sort of risky behavior may depend on his or her beliefs regarding how many peers are also engaging in that behavior. A complete technical overview of interaction-based models is beyond the scope of this report, but one point worth noting is that the more that individuals' decisions are influenced by the decisions of others (if this influence is positive, it would imply a conformity effect), the greater the likelihood that the social dynamics implied by the process have multiple possible stable outcomes (i.e., equilibria). See pages 66–67 in Durlauf (2001) for a discussion of this issue.

This framework for capturing interdependent decisions has been applied to studies of peer effects on smoking. For example, Card and Giuliano (2013) use information on friendship ties from the National Longitudinal Study of Adolescent Health (Add Health) to estimate discrete choice models of adolescents' choices concerning smoking, sex, and truancy. The researchers find some evidence of social interactions, especially with regard to peer effects on sexual activity. For example, having a best friend who is sexually active increases the likelihood that one is sexually active by 5 percentage points. Weaker evidence also supports peer effects with regard to smoking, marijuana, and truancy.

Qualitative Approaches

Another strategy for modeling decision making is to specify a heuristic rule based on experimental or theoretical knowledge of the process to be modeled and to assume that agents in the model use that rule to make decisions. Heuristics are "rules of thumb" for making decisions under conditions of uncertainty (Kahneman et al., 1982). Heuristics can be invoked both when gathering information to inform decision making and when evaluating information in the actual decision. Heuristics may be combined with a set of weights that specify the relative desirability of various alternative choices. For example, once a set of choice options has been evaluated, one must decide how to go about choosing among them. One option is to use a "satisficing" heuristic—that is, to assume that people are indifferent among various alternative choices as long as they all satisfy some baseline level of acceptability. In the absence of hard evidence about how people go about making decisions, the decision-making mechanism is yet another assumption that goes into model specification. One fruitful area for future research would be pinning down how real people make decisions.

Data may be used to specify agents' behavior by using ethnographic or participant observations that can provide information on the motivations, strategies, or "rules of thumb" that drive decision making. For example, Hoffer et al. (2009) use ethnographic data to calibrate their ABM of heroin markets. In contrast to statistical specifications of behavior, a qualitative model of behavior is typically formulated as a set of rules governing human action or, alternatively, as a set of rules for interpreting information. One can also combine quantitative and qualitative data on behavior. For example, if experiments reveal a systematic bias in how people perceive their environment, an adjustment could be made to the inputs of a statistical model of behavior.

MODEL UNCERTAINTY AND POLICY DECISION MAKING

As should be obvious from the discussion thus far, models cannot predict the future with certainty. Models can mislead policy makers if modelers present their findings with greater certitude than is warranted. A good model will quantify how uncertainty in the model's inputs translates into uncertainty in what outcomes are most likely under a given policy and will generate a range of predictions that reflect that uncertainty (Manski, 2013;

Wagner et al., 2010). The key issue is separating what is known from what is unknown. Note that this is a very different enterprise from conducting a "parameter sweep" type of sensitivity analysis, which merely provides insight into the workings of the model itself and not into the relationship between the model and the actual world. Uncertainty is only meaningful if the model is anchored in key features of the process under investigation. At a minimum, this might be a simple model that includes an empirically defensible representation of individuals' behavior and interaction.

Once analysts have generated a set of credible model outputs, they must use that information to draw some kind of conclusion about the best course of action. The challenge for the policy maker is to evaluate candidate policy outcomes and weigh the risks and benefits. Thus, to use models effectively to guide policy decisions, the model user needs a rule for translating uncertain predictions into a policy decision.

Conclusion 3-4: The committee concludes that the common exercise of sensitivity analysis does not suffice to measure the uncertainty in model-based forecasts. Sensitivity analysis may provide some insight into the workings of the model itself, but it does not per se assess the potential relationship between model findings and the real world.

Recommendation 3-3: When the U.S. Food and Drug Administration uses the findings of any model, the agency should take into account the uncertainty of findings in order to evaluate policy outcomes and weigh the risks and benefits appropriately.

CONCLUSION

In this chapter, the committee provides an overview of the use of ABMs in policy decision making and explicates how ABM fits into a larger set of modeling approaches. The committee found that ABMs could play an important role in policy decision making and offer useful insights that are not possible with a more aggregated approach. However, to provide meaningful inferences, ABMs must at a minimum include a plausible representation of individual behavior. This may be a fruitful avenue for future research. Moreover, models must provide some account of how uncertainty in model inputs translates into uncertainty in model outputs. To use these models effectively, policy makers will likely need to develop a rule for translating these uncertain predictions into a policy decision.

REFERENCES

Allison, P. D. 1982. Discrete-time methods for the analysis of event histories. *Sociological Methodology* 13(1):61–98.

Aral, S., L. Muchnik, and A. Sundararajan. 2009. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences of the United States of America* 106(51):21544–21549.

Arno, P. S., A. M. Brandt, L. O. Gostin, and J. Morgan. 1996. Tobacco industry strategies to oppose federal regulation. *JAMA* 275(16):1258–1262.

Ben-Akiva, M. E., and S. R. Lerman. 1985. *Discrete choice analysis: Theory and application to travel demand*. Vol. 9. Cambridge, MA: MIT Press.

Blume, L. 2015. Agent-based models for policy analysis. Paper commissioned by the Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation (see Appendix B).

Blume, L. E., W. A. Brock, S. N. Durlauf, and Y. M. Ioannides. 2010. Identification of social interactions. In *Handbook of social economics*, Vol. 1b, edited by J. Benhabib, A. Bisin, and M. Jackson. Amsterdam, Netherlands: North Holland, Elsevier.

Brock, W. A., and S. N. Durlauf. 2001. Discrete choice with social interactions. *The Review of Economic Studies* 68(2):235–260.

Brown, D. G., S. E. Page, R. Riolo, and W. Rand. 2004. Agent-based and analytical modeling to evaluate the effectiveness of greenbelts. *Environmental Modelling & Software* 19(12):1097–1109.

Bruch, E., and J. Atwell. 2013. Agent-based models in empirical social research. *Sociological Methods & Research* 1–36.

Card, D., and L. Giuliano. 2013. Peer effects and multiple equilibria in the risky behavior of friends. *Review of Economics and Statistics* 95(4):1130–1149.

Cartwright, N. 2007. Are RCTs the gold standard? *BioSocieties* 2(1):11–20.

CBO (Congressional Budget Office). 2007. *The effect of tax changes on labor supply in CBO's microsimulation tax model*. Washington, DC: Congressional Budget Office.

Chaloupka, F., and R. Pacula. 1999. The impact of price on youth tobacco use. In *Changing adolescent smoking prevalence: Where it is and why. Smoking and Tobacco Control Monograph 14*. Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health, National Cancer Institute.

Chaloupka, F. J., and K. E. Warner. 2000. The economics of smoking. In *Handbook of health economics edition 1*. Vol. 1, edited by A. J. Culyer and J. P. Newhouse. Amsterdam, Netherlands: North Holland, Elsevier. Pp. 1539–1627.

Christakis N. A., and J. H. Fowler. 2008. The collective dynamics of smoking in a large social network. *The New England Journal of Medicine* 358(21):2249–2258.

Devereux, M. P., and G. Maffini. 2007. *The impact of taxation on the location of capital, firms and profit: A survey of empirical evidence*. Oxford University Centre for Business Taxation. https://ideas.repec.org/p/btx/wpaper/0702.html#biblio (accessed March 18, 2015).

Durlauf, S. N. 2001. A framework for the study of individual behavior and social interactions. *Sociological Methodology* 31(1):47–87.

Epstein, J. M., J. Parker, D. Cummings, and R. A. Hammond. 2008. Coupled contagion dynamics of fear and disease: Mathematical and computational explorations. *PLoS ONE* 3(12):e3955.

Fletcher, J. M. 2010. Social interactions and smoking: Evidence using multiple student cohorts, instrumental variables, and school fixed effects. *Health Economics* 19(4):466–484.

Fletcher, J. M., and S. L. Ross. 2012. *Estimating the effects of friendship networks on health behaviors of adolescents*.Cambridge, MA: National Bureau of Economic Research.

Friedman, M. 1953. *Essays in positive economics*. Chicago, IL: University of Chicago Press.

Gilbride, T. J., and G. M. Allenby. 2004. A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Science* 23(3):391–406.

Grazzini, J., M. G. Richiardi, and L. Sella. 2013. *Analysis of agent-based models*. Laboratorio R. Revelli, Centre for Employment Studies. http://www.laboratoriorevelli.it/_pdf/wp135.pdf (accessed March 18, 2015).

Hammond, R. A. 2015. Considerations and best practices in agent-based modeling to inform policy. Paper commissioned by the Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation (see Appendix A).

Hauser, J. R., O. Toubia, T. Evgeniou, R. Befurt, and D. Dzyabura. 2010. Disjunctions of conjunctions, cognitive simplicity, and consideration sets. *Journal of Marketing Research* 47(3):485–496.

Heckman, J. J. 2008. Econometric causality. *International Statistical Review* 76(1):1–27.

Henriksen, L. 2012. Comprehensive tobacco marketing restrictions: Promotion, packaging, price and place. *Tobacco Control* 21(2):147–153.

HHS (U.S. Department of Health and Human Services). 2010. *How tobacco smoke causes disease: The biology and behavioral basis for smoking-attributable disease: A report of the Surgeon General.* Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention, Office on Smoking and Health.

Hoffer, L. D., G. Bobashev, and R. J. Morris. 2009. Researching a local heroin market as a complex adaptive system. *American Journal of Community Psychology* 44(3-4):273–286.

IOM (Institute of Medicine). 2012. *Scientific standards for studies on modified risk tobacco products*. Washington, DC: The National Academies Press.

Jeon, J., R. Meza, M. Krapcho, L. D. Clarke, J. Byrne, and D. T. Levy. 2012. Chapter 5: Actual and counterfactual smoking prevalence rates in the U.S. population via microsimulation. *Risk Analysis* 32(Suppl 1):S51–S68.

Kahneman, D., P. Slovic, and A. Tversky. 1982. *Judgment under uncertainty: Heuristics and biases*. Cambridge, MA: Cambridge University Press.

Katz, M. H., S. K. Schwarcz, T. A. Kellogg, J. D. Klausner, J. W. Dilley, S. Gibson, and W. McFarland. 2002. Impact of highly active antiretroviral treatment on HIV seroincidence among men who have sex with men: San Francisco. *American Journal of Public Health* 92(3):388–394.

Lightfoot, M., D. Swendeman, M. J. Rotheram-Borus, W. S. Comulada, and R. Weiss. 2005. Risk behaviors of youth living with HIV: Pre-and post-HAART. *American Journal of Health Behavior* 29(2):162–171.

Loomis, B. R., M. C. Farrelly, and N. H. Mann. 2006. The association of retail promotions for cigarettes with the Master Settlement Agreement, tobacco control programmes and cigarette excise taxes. *Tobacco Control* 15(6):458–463.

Macy, M. W., and R. Willer. 2002. From factors to actors: Computational sociology and agent-based modeling. *Annual Review of Sociology* 28(1):143–166.

Manski, C. F. 1993. Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies* 60(3):531–542.

———. 2000. Economic analysis of social interactions. *Journal of Economic Perspectives* 14(3):115–136.

———. 2007. *Identification for prediction and decision*. Cambridge, MA: Harvard University Press.

———. 2013. *Public policy in an uncertain world:Analysis and decisions*. Cambridge, MA: Harvard University Press.

Mare, R. D., and E. E. Bruch. 2003. Spatial inequality, neighborhood mobility, and residential segregation. On-Line Working Paper Series, California Center for Population Research, University of California, Los Angeles. http://escholarship.org/uc/item/8xm4q2q7 (accessed March 18, 2015).

Marschak, J. 1974. Economic measurements for policy and prediction. In *Economic information, decision, and prediction*, Vol. 7–3. Netherlands: Springer. Pp. 293–322.

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. Chapter 4. In *Frontiers in econometrics*, edited by P. Zarembka. New York: New York Academic Press. Pp. 105–142.

Merton, R. K. 1936. The unanticipated consequences of purposive social action. *American Sociological Review* 1(6):894–904.

Moffitt, R. A. 2001. Policy interventions, low-level equilibria, and social interactions. Chapter 3. In *Social dynamics*, edited by S. N. Durlauf and H. P. Young. Washington, DC: Brookings Institution Press. Pp. 45–82.

NCI (National Cancer Institute). 2001. *Risks associated with smoking cigarettes with low machine-measured yields of tar and nicotine*. NIH publication No. 02-5074. Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health, National Cancer Institute.

NRC (National Research Council). 2007. *Models in environmental regulatory decision making*. Washington, DC: The National Academies Press.

———. 2012. *Assessing the reliability of complex models: Mathematical and statistical foundations of verification, validation, and uncertainty quantification*. Washington, DC: The National Academies Press.

———. 2014. *Advancing land change modeling: Opportunities and research requirements*. Washington, DC: The National Academies Press.

Osgood, N. 2007. Using traditional and agent based toolsets for system dynamics: Present tradeoffs and future evolution. Presented at International Conference of the System Dynamics Society, Boston, MA. http://www.systemdynamics.org/conferences/2007/proceed/papers/OSGOO488.pdf (accessed March 18, 2015).

Saez, E., J. Slemrod, and S. H. Giertz. 2012. The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature* 50(1):3–50.

Salop, S., and J. E. Stiglitz. 1982. The theory of sales: A simple model of equilibrium price dispersion with identical agents. *The American Economic Review* 72(5):1121–1130.

Sampson, R. J., C. Winship, and C. Knight. 2013. Overview of: "Translating causal claims: Principles and strategies for policy-relevant criminology." *Criminology & Public Policy* 12(4):585–586.

Sanstad, A. H. 2015. Assessing agent-based models for regulatory applications: Lessons from energy analysis. Paper commissioned by the Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation (see Appendix C).

Schelling, T. 1978. *Micromotives and macrobehavior*. New York: Norton.

Shalizi, C. R., and A. C. Thomas. 2011. Homophily and contagion are generically confounded in observational social network studies. *Sociological Methods & Research* 40(2):211–239.

Smith, A. 1759. *The theory of moral sentiments*. Edited by D. D. Raphael and A. L. Macfied. Oxford, UK: Oxford University Press.

Sterman, J. D. 2006. Learning from evidence in a complex world. *American Journal of Public Health* 96(3):505–514.

van Meijgaard, J., J. E. Fielding, and G. F. Kominski. 2009. Assessing and forecasting population health: Integrating knowledge and beliefs in a comprehensive framework. *Public Health Reports* 124(6):778–789.

Wagner, W., E. Fisher, and P. Pascual. 2010. Misunderstanding models in environmental and public health regulation. *NYU Law Environmental Law Journal* 18:293–356.

Wilde, G. J. S. 2001. *Target risk 2: A new psychology of safety and health: What works? What doesn't? And why…* Toronto, Canada: PDE Publications.

4

An Evaluation Framework for Policy-Relevant Agent-Based Models

Policy-relevant agent-based models (ABMs) are resource-intensive, complex technical activities that are developed by large groups of people with varying areas of expertise. The results of these models need to be translated and communicated to various stakeholders in order to affect policy and improve health. Policy-relevant ABMs need to be built carefully using appropriate data and social science theories, rigorously tested, and clearly communicated. These requirements for ABMs are the same as for other types of computational models and simulations used to inform policy decisions (e.g., aggregate models, system dynamics, and econometric forecasting models, to name a few).

Given the amount of time, effort, and money required to build an effective policy-relevant model, it is critical to evaluate the process, its outcomes, and its overall value. Was the model built such that its results represent what the modelers intended? Did the model address important and timely policy questions? Were the results useful in guiding subsequent policy and regulatory decision making and research? And, in the end, were the results worth the cost? These types of questions can be answered by *evaluating* the model building process, the model outcomes, and the return on investment.

The goal of this chapter is to present an evaluation framework for assessing ABMs for tobacco control policy and regulation. The committee found that no such framework exists for tobacco control and that such a framework is needed to assess complex computational modeling projects in a wide variety of public health policy and regulatory contexts. The committee developed this evaluation framework both to guide the committee in its review of the model developed for the U.S. Food and Drug Administration

(FDA) (see Chapter 5), and to provide FDA with guidance for future model development and evaluation.

An evaluation framework for ABMs can provide answers to two broad types of questions:

a. *Process*—How was the model informed (by subject-matter experts, by data, and other inputs), planned, developed, and tested?
b. *Outcomes*—In what ways did the modeling produce results that were useful for guiding future policy and regulatory efforts?

The remainder of this chapter provides a grammar for describing policy-relevant ABMs, presents an evaluation framework, identifies high-priority categories for evaluation questions, and illustrates some of the evaluation concepts through two case-study descriptions of existing policy-relevant ABMs.

MODEL DESCRIPTION

Before establishing a framework for evaluating how ABMs inform public health or tobacco control policy and regulation, a consistent way to describe and talk about them is needed. Although there have been some attempts at classifying ABMs to aid in model description (e.g., Marietto et al., 2003), these have tended to be too broad to capture the diversity of types of models that can be helpful for advancing policy and regulation or else too technical and not applicable to policy-relevant models (e.g., Grimm et al., 2006).

Table 4-1 presents a set of model descriptors that can be thought of as a *grammar* for describing in detail the structure and purpose of a policy-relevant ABM. This set of descriptors is not meant to replace a complete technical description of the model (sometimes called a "design document"). Instead, this gives a formal way to provide a rich description of the important elements of the model to be evaluated. Model evaluation requires a concise but thorough description of the model and what it was designed to accomplish.

The descriptors in Table 4-1 fall into seven broad categories: basic model description, model agents, use of data and theories, model context, model outcomes, policy aspects, and communication aspects. Within each of these categories is a small set of individual descriptors. For example, *physical space* is the indicator under *context* that describes whether and how the ABM depicts the physical space within which agents are allowed to move. Consistent use and application of these terms during model development will lead to better communication among the model developers and users of the model and will maximize the chance that the model meets the

TABLE 4-1 Grammar for Describing Agent-Based Models Relevant for Policy and Regulation

Descriptors	Definition
	Basic Description
Purpose (goal)	What is the main scientific, policy, or regulatory question that the model is addressing?
Breadth	What is the scope of the model? Is it designed to focus narrowly on one or a small number of social system components or processes, or is it designed to broadly encompass most or all parts of a complex system?
Abstraction	Is the model designed to be highly abstract, with the agents, rules, and context (i.e., physical and social spaces) not meant to precisely match real world settings, behavior, and processes, or is it designed with realism in mind?
	Agents
Agent type	Does the model include one type of agent, or multiple types (i.e., a multi-agent model)?
Agent definition	What are the agents in the model? For example, are the agents people or some other type of social agent (e.g., tobacco retailer)?
	Data and Theories
Data—rules	Are empirical data (quantitative or qualitative) used to inform the agent rules (e.g., smoking prevalence used to shape smoking initiation decision by an agent)?
Data—characteristics	Are empirical data used to inform the characteristics of the agents and environment?
Data—validation	Are empirical data used to validate model results?
Theories	What are the primary social science and behavioral theories used in the model design?
	Context
Physical space	Does the model include an explicit depiction of the physical space (e.g., built environment, geography) within which agents are allowed to move?
Social space	Does the model include an explicit depiction of the social space (i.e., connections or relationships among social entities such as people, communities, and organizations) that influences agent behavior or structures flow of information or other resources?
Physical dynamics	Is the physical space static or allowed to change as part of the model?
Social dynamics	Is the social space static or allowed to change as part of the model?

continued

TABLE 4-1 Continued

Descriptors	Definition
	Outcomes
Primary outcome	What is the primary outcome that is being modeled?
Proximal/distal outcome	Is the primary outcome a proximal or distal behavioral indicator? For example, reduction of smoking prevalence may be the ultimate goal of a policy that is being modeled, but the model may focus on addictive properties of new products or new restrictions on marketing. In these cases, these would be considered proximal outcomes.
	Policy
Policy definition	Description of the policy or policies that are being examined in the models.
Policy realism	How realistic are the policies being examined in the model? Are they reflective of actual policies that are being implemented, or do they reflect more abstract policy mechanisms or classes?
Policy tests	Does the model include formal tests of policy effects?
	Communications
Model sharing	What aspects of the model are (or will be) publicly available?

NOTE: The grammar in this table is meant to offer guidance on how to describe an ABM and is not meant to provide an evaluation of the quality of the model, which is something that is done later in the model development process.

needs of the model sponsor (Kuntz et al., 2013). See the chapter annex (see Table 4-2) for examples of how the descriptive grammar can be applied to three different policy-relevant ABMs.

EVALUATION FRAMEWORK

Fundamentally, systematic evaluations of policy-relevant ABMs are important because they can lead to better and more effective models in the future. A comprehensive evaluation provides useful information to at least four different groups involved with models:

a. It helps the model *developers* improve their modeling efforts;
b. It helps the *funders* understand better how to use model results and how to guide future funding of modeling work;
c. It helps *policy makers* understand how to translate model results into more effective policies and increases their trust in the analysis; and
d. It helps *modelers* and *scientists* by suggesting new avenues for research, modeling, and data collection.

Figure 4-1 presents an evaluation framework for policy-relevant ABMs that can be used specifically to evaluate models designed to inform tobacco control policy and regulation. The framework uses a logic-model approach, following the U.S. Centers for Disease Control and Prevention's program evaluation framework (CDC, 1999, 2007), and is based on best practices identified in a number of modeling fields. Although logic models have been used primarily to guide program evaluation, they are also useful for designing evaluations of policy development and implementation (Jordan, 2010; Langer et al., 2011) and larger systems evaluations (CDC, 2011; CORE, 2009). A logic model helps to ensure that all important aspects of the model building process are included in a systematic evaluation, and it guides the prioritization of evaluation questions. Logic models define the domains that are important for understanding the relevant processes and outcomes; however, actual evaluations based on the logic model will typically focus on a subset of domains (and associated evaluation questions) (CDC, 2007). The evaluation framework is not meant to be used as a checklist—each area deemed relevant to a particular model requires consideration on how each decision point will affect the model in the end.

The development of the evaluation framework was based on a review of relevant literature and on committee members' experience in building and assessing ABMs and in developing public health policy evaluations. The evaluation framework is designed to cover the important aspects of designing, implementing, testing, and disseminating policy-relevant ABMs, especially for tobacco control regulatory and policy efforts. It can be used to assess the model development processes as well as its outcomes. The framework has five major sections—resources, activities, outputs, outcomes, and environment—any of which can be the focus of a systematic evaluation.

Model development is an iterative process, and there are no clear divisions between the various evaluation steps (Berk, 2008), so there will be some overlap between the domains of the major sections (resources, activities, outputs, outcomes, and environment) of the logic model. It is often the case that a logic model will have boxes with the same or similar names across the columns (CDC, 2011). For example, there are policy activities that lead to policy outputs, which in turn influence policy outcomes. One important reason for using this structure is that the evaluation questions and timeline are quite different for early activities versus long-term outcomes.

Resources

A successful policy-relevant ABM is made up of a wide variety of critical ingredients. The domains listed in the resources section of the framework reflect the most important of these individual elements, which are the people, knowledge, infrastructure, and financial support neces-

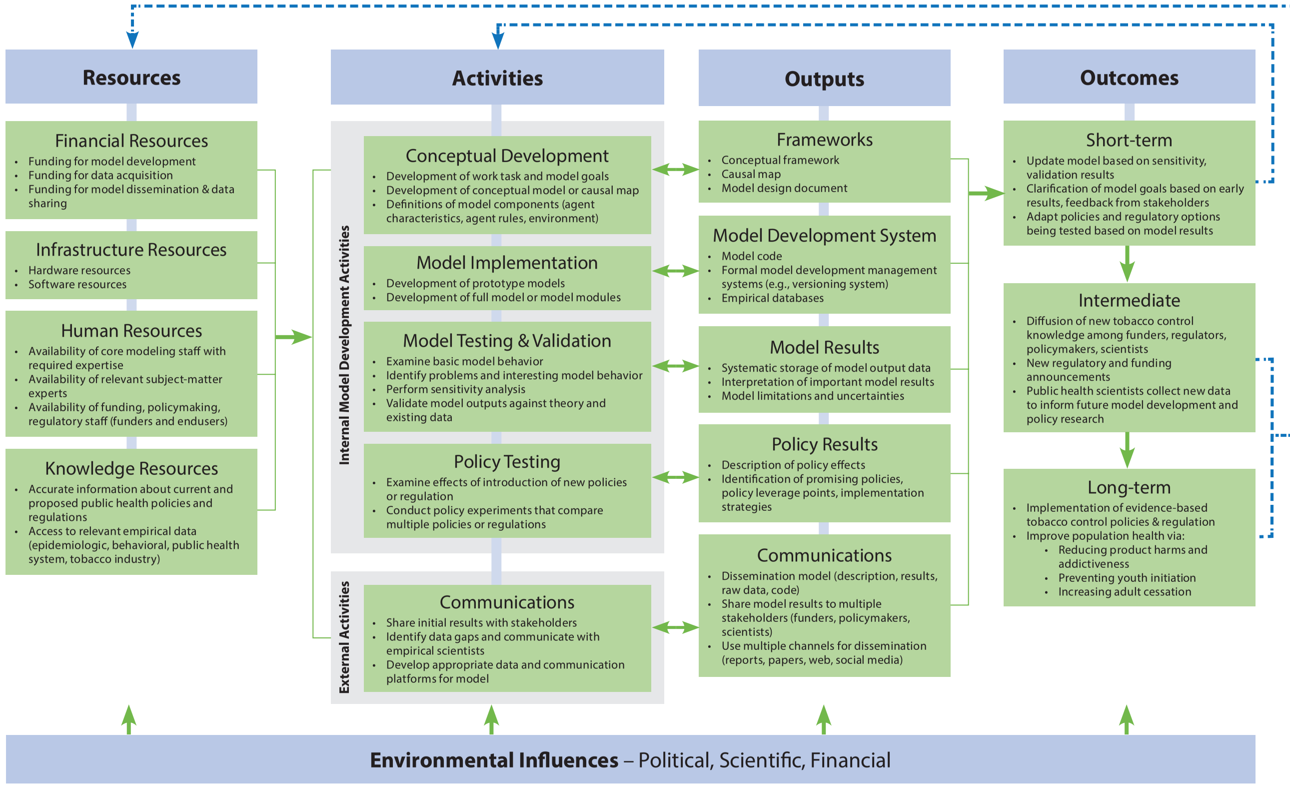

FIGURE 4-1 An evaluation framework for policy-relevant agent-based models.

sary for the successful development of an ABM. Typically, for example, a sophisticated modeling effort would not be started without the necessary funding and infrastructure already in place, but additional resources are often needed for data acquisition, model alterations, or an expanded dissemination of model findings, without any of which the model effectiveness could be diminished. Even more important for the successful development of a model are the necessary people and knowledge. Before model development begins, discussions with modelers from varied backgrounds (such as economics, engineering, and social sciences, among others) can help identify the best modeling approach (or combination of approaches) for the policy question under consideration (Roberts et al., 2012). These conversations can help ensure that an appropriate modeling strategy is chosen from the outset. In addition to the core modeling staff, a model must be informed by input from relevant subject-matter experts (Kuntz et al., 2013). For tobacco control policy and regulatory models, these may include experts from the social, clinical, and basic sciences. Another important group of people consists of the model funders and other policy makers who may use the results of the model to inform their policy and regulatory work. Finally, a modeling team will need access to a wide variety of knowledge resources, including relevant data, empirical findings, and current and proposed policies and regulatory options that may be addressed by the ABM.

Activities

The actions required to develop a policy-relevant ABM are listed under the activities section of the logic model. The first three boxes contain those activities that are common to any ABM, or almost any type of computational model. Model development starts with a conceptual phase, followed by implementing the model in code and then performing a series of validation tests. It is particularly important that the initial conceptual phase is included in an evaluation, yet this does not always happen:

> Some of the most important model choices are made at the conceptual stage, yet most model evaluation activities tend to avoid a critical evaluation at this stage. Often a peer review panel will begin its efforts with the implicit acceptance of all the key assumptions made to establish the conceptual model and then devote all of its attention to the model building and model application stages. (NRC, 2007, p. 115)

The technical requirements for strong model development are many and varied, but the modeling community has developed a number of systematic approaches for describing, managing, and monitoring this development (Badham, 2010; Berk, 2008; Caro et al., 2012; CREM, 2009; Gurcan et al., 2011; Helbing and Balietti, 2011; NRC, 2007, 2012; Šalamon, 2011;

Weinstein et al., 2003).[1] (See Appendix A for a broad overview of some of these requirements as well as the PARTE framework, which provides a helpful framework for defining model agents and context.) Some of these technical assessments can be useful for an evaluation of the model development process, but it is important to distinguish a more technical assessment of model validation, verification, and uncertainty quantification from a more general evaluation of the modeling processes and model outcomes. As also noted in a 2007 report by the National Research Council (NRC), the process of model evaluation is more than just a strict validation or verification procedure; it is a process that builds confidence in model applications and increases the understanding of model strengths and limitations. "Regulatory model evaluation must consider how accurately a particular model application represents the system of interest while being reproducible, transparent, and useful for the regulatory decision at hand" (NRC, 2007, p. 3). In essence, models should be evaluated with regard to their suitability as tools to address the question or process under study.

Iterative data collection throughout the model development process is often crucial in the development of effective policy-relevant models. In many cases, it can be difficult to predict what data may be needed to parameterize the model before—and even during—model development. When the model is implemented and outputs are generated, data gaps may be identified, signifying areas where data (whether available or not) is needed to inform a critical component of the model. Having identified these gaps, the model developers may revamp their model with newly integrated available data, or they may acknowledge the limitations of the data and encourage further data collection so that future models addressing a similar purpose can be more useful for the regulatory decision at hand. (See Case Examples later in this chapter for an illustration of this point.) Close and ongoing communication with subject-matter experts (detailed on the next page) can facilitate identifying these gaps early in the process.

Not all ABMs are intended to test the effects of different policies, but such policy testing is the raison d'être for policy-relevant models. Develop-

[1] For example, Berk provides a six-stage evaluation process based on the work of Bayarri and colleagues (2007). These six stages are model specification, including the interactions between agents; the determination of model evaluation criteria, including calibration (fit) and use of visualizations; data collection of model inputs (for calibration) and "ground truth" test data (for model outputs), which refers to any data that capture the empirical process under investigation; construction of model approximations, e.g., statistical summaries of inputs and outputs and data reduction; an analysis of model output, e.g., search for obvious discrepancies between ground truth and model output; and the feedback of information to model development, while avoiding turning test data into calibration data.

ing informative policy-testing ABMs is challenging,[2] but once such an ABM has been developed, it can be used to explore the hypothesized effects of specific policy or regulatory interventions to reveal the possible mechanisms by which these policies operate or even to perform in silico policy experiments where different policy options are compared to one another (Auchincloss and Diez Roux, 2008; Hammond, 2015; Lempert, 2002). Finally, throughout all model building phases, it is necessary to perform a number of communication tasks. Funders and sponsors need to be kept apprised of progress, and content experts need to talk frequently with the model building staff to avoid making major errors related to essential model implementation decisions (Kuntz et al., 2013), such as programming agent behaviors, and to identify data gaps throughout model development. Because complex problems such as tobacco control require collaborative, interdisciplinary efforts—and thus varying backgrounds among the model-building stakeholders—communication will inevitably be interdisciplinary in nature. It can be a difficult process to get all team members on the same page early in the modeling process; there can be differences in the use of terminology, approaches, and strategy (Hovelynck et al., 2010; Nicolson et al., 2002). However, this process results in a stronger model in the end. It is often helpful to have as part of the team a translator or "knowledge broker" (Bammer, 2012; Bammer et al., 2010; Bielak et al., 2008; Dobbins et al., 2004; Meyer, 2010) who has a solid understanding both of the policy issues at hand and of modeling (while likely not actually being a modeler) and who can ensure that the information from the subject-matter experts is effectively translated into the model.

Outputs

Each of the model building activity domains has an associated set of products and outputs. For example, the outputs from the conceptual development stage may include causal maps, conceptual frameworks, and the general model design document. These outputs are often the primary subjects of process-focused evaluations. For example, if a model evaluation is focused on the validity of the agent behavioral rules incorporated in the model, the model design document will be an important source of information on how the agent behaviors were constructed. To ensure that the end users understand the scope of the model and properly apply the modeling results during development and upon completion of the model, the documentation of model's limitations and uncertainties is an important output (Eddy et al., 2012).

[2]See Appendixes B and C for a comprehensive discussion on the challenges of developing informative policy-testing ABMs.

Outcomes

A variety of short- and longer-term outcomes is typically of interest when evaluating a policy-relevant ABM. The short-term outcomes include the immediate results of the model both as it is being developed and right after the model is finalized. The longer-term effects of a model include the diffusion of its results across a variety of stakeholder audiences (e.g., regulators, policy scientists, policy makers, tobacco content experts, and other modelers). Ultimately, models may result in changes in the policy and regulatory environments, shifts in funding priorities, changes in the types of data collected, implementation of new policies and regulations, and subsequent changes in the behavior of individuals, of the public health sector, and of organizations (e.g., tobacco companies). Although these longer-term effects are of obvious interest, by their very nature they take a long time to manifest themselves. In addition, whether model results are used by policy makers is largely out of the control of the model development team. Even when a model has useful outputs that are effectively quantified and communicated, the policy maker might not understand the value that the model has to offer (Kuntz et al., 2013), or unexpected changes in the environment can make the model results outdated (NRC, 2007). Thus, even in an outcome-focused ABM evaluation, long-term policy outcomes may not be explicitly included. However, this highlights the need for the policy maker to be involved with model development from its conception and the importance of translating the model results properly. Although this involvement will not guarantee that the model is used, it will increase the likelihood that the model addresses the current questions the policy makers are faced with and that they have a deep understanding of the value that the model offers (Wagner et al., 2010).

Environment

A variety of external environmental characteristics and forces might influence model development, either positively or negatively. For example, a shift in governmental policy priorities may make certain modeling efforts of greater interest to stakeholders. Alternatively, a change to federal, state, or local tobacco laws might affect tobacco use patterns, which would then need to be accounted for in a model. The rapid introduction of non-combustible tobacco products (e.g., e-cigarettes) in recent years, which is having dramatic effects on tobacco product purchasing and consumption as well as on industry and retailer behavior, is a good example of the sort of environmental change that can influence policy-relevant model development. Environmental factors are typically not the focus of an evaluation, but the entity conducting the evaluation needs to be aware of these influences so that the evaluation results and conclusions can be put into the appropriate context (Weinstein et al., 2001).

In addition to the individual domains listed in the logic model, the framework also focuses on some of the relationships between the domains, which illustrate the dynamic nature of ABM development. In particular, three types of feedback loops are present throughout model development and dissemination. First, the direct outputs of model development (i.e., conceptual frameworks and causal maps, model code, model testing and validation results, and policy testing results) provide feedback directly to the modeling team. In particular, the validation results invariably lead to modeling changes and improvements. (This feedback is depicted via double-headed arrows connecting the activities and outputs boxes.) Second, modeling results are typically disseminated in a number of ways, including via meetings with funders, reports, conference presentations, and peer-reviewed scientific papers. Immediate reactions to this dissemination can lead to further data collection, model development, or model expansion. For example, funders may ask the modeling team to consider new types of policy experiments or questions based on initial model results. Or new data may be made available that could be used to improve or expand the initial model. (This is depicted by the inner feedback loop connecting the short-term outcomes box to model activities.) Finally, intermediate and long-term outcomes constitute the types of policy, regulatory, and public health changes that were the goals of the modeling in the first place. These major changes in the policy and health landscapes will lead to completely new modeling efforts. (This feedback is depicted by the outer line connecting the intermediate and long-term outcomes boxes to future model resources.) And, of course, model development does not end here, as it is an iterative process. In particular, it is useful to think about taking a life-cycle approach to model development and testing (NRC, 2007).

Identifying High-Priority Evaluation Questions

The evaluation framework for policy-relevant ABMs provides a guide for designing an evaluation of a specific ABM project or a broader modeling initiative. Evaluations can be used to answer many questions, but they are most effective when there is a clearly stated purpose for the evaluation. For example, the main purpose of an ABM outcome evaluation may be to identify how the modeling results influenced subsequent policy and regulatory research. Once this overall purpose is decided on, the next task is to identify the set of specific questions that will be addressed in the evaluation. Despite the broad nature of the framework presented in this chapter, it would not be feasible to have an evaluation focus on every single domain. Instead, a short list of specific evaluation questions should be identified that are most important to implement and to derive from the overall purpose of the evaluation. Typically, this prioritization process starts with a longer list of potential questions that can be linked to the logic model, and the list is then shortened by deciding which questions are most important. In this chapter's annex, the committee provides a list of example questions based on the evaluation framework that could be appropriate for evaluating an ABM project, especially in the context of tobacco control policy and regulatory science.

Fundamental Evaluation Categories

As the committee developed its evaluation framework, five fundamental evaluation categories for policy-relevant ABMs emerged. These are broader categories of relevant evaluation and assessment domains that the committee believes need to be included in most policy-relevant ABM evaluations. The five categories are listed below, as well as some sample questions for consideration (with a longer list of evaluation questions available in the chapter annex).

a. **Resources:** A modeling team needs access to adequate financial, infrastructure, human, and knowledge resources to successfully design, build, and test its model.
   - To what extent were relevant staff available (e.g., funders, policy makers, end users) as the model was being built, especially in the conceptual development phase?

b. **Technical best practices:** Model implementation, testing, and validation phases need to be reviewed throughout model development.
   - What kinds of analyses were done to quantify uncertainty?
   - How do the results compare to the results of other models addressing similar policy questions or having similar purpose?

c. **Model suitability:** Models need to be developed in a manner that makes them suitable for their intended purpose and that will allow for exploration or testing of specific policy options or conditions. Some models could be developed for answering very narrow questions related to tobacco use, others as broader tools to look at a larger range of tobacco policies.
   - Does the goal of the model match the methods used and the assumptions made?
   - To what extent does the model capture the fundamental dynamics thought to be operating in the real world?
   - Does the model provide information that is helpful to making tobacco control policy?

d. **Communication and translation:** Communication and translation strategies are essential during every stage of model development for enhancing the model building process and ensuring that the model is focused on the key issues that will affect policy outcomes. Modeling requirements, descriptions, and results need to be communicated effectively to a variety of audiences, including agency staff, regulators, politicians, and the general public.
   - Does the model documentation include a write-up of model uncertainties, interpretations of results, and considerations for maintenance of the model?
   - How were preliminary results fed back into subsequent model improvements?
   - Were model processes and results communicated in a manner that allows for reproducibility?
   - If proprietary issues and requirements limited the communication of modeling information, were the costs and benefits of those limitations assessed or articulated?

e. **Policy outcomes:** Ultimately, policy-relevant models will be used to inform policy and regulatory action or to advance scientific progress. Many of the likely evaluation questions in this category are not in the control of the model development team, and policy-related evaluation results do not necessarily reflect the quality of the model, but this reinforces the need for collaboration with policy makers from the onset.
   - Was the model used to inform policy decisions? Did policies and regulatory options change in response to the model results?
   - How flexible is the model (i.e., capacity for the model to be modified or revised and applied to situations as new data arise or alternative objectives are specified)? What factors might trigger the need for major revisions, or what circumstances might prompt users to seek an alternative model?
   - How has the sponsor (e.g., FDA) used model results to inform its own regulatory activities?
   - How relevant are the modeling results to the tobacco control field? Have the results informed tobacco control knowledge and influenced decisions among funders, regulators, policy makers, scientists?

Recommendation 4-1: The Center for Tobacco Products should adopt an evaluation framework for its modeling work, either the one presented in this chapter or one similar to it. Key dimensions of the evaluation framework should include considerations of resources, technical best practices, model suitability, communication and translation, and

policy outcomes. The evaluation plan should be designed early in the model development process and should be carried out throughout model development.

This evaluation framework would apply to all efforts funded by the Center for Tobacco Products (CTP)—internal model development, interagency agreements, contracts, and grants. In addition to internal CTP reviewers, external experts need to be part of the evaluation process (see NRC, 2007, for guidance on the peer review of models).[3] If CTP chooses to adopt the evaluation framework developed by the committee, the framework should be used as a guideline and not as a mechanical exercise or checklist, because different ABMs will require differing evaluation strategies based on intended use, modeling approach, and other aspects of model development.

CASE EXAMPLES

In this section, the committee explores published models that illustrate many of the areas outlined in the evaluation framework. These examples cover subjects from two different areas: transportation and illicit drugs. The committee chose these examples because they illustrate several of the important aspects of model development discussed in this chapter; however, the committee did not formally review or assess the overall strengths or weaknesses of the models. It is difficult to provide examples of all of the

[3] Guidance on peer review can be found in NRC, 2007. Options for receiving external review include contracts, special government appointees, and advisory panels.

> Peer review should be considered, but not necessarily performed, at each stage in a model's life cycle. Some simple, uncontroversial models might not require any peer review, whereas others might merit peer review at several stages. Appropriate peer review requires an effort commensurate with the complexity and significance of the model application. When a model peer review is undertaken, EPA should allow sufficient time, resources, and structure to assure an adequate review. Reviewers should receive not only copies of the model and its documentation but also documentation of its origin and history. Peer review for some regulatory models should involve comparing the model results with known test cases, reviewing the model code and documentation, and running the model for several types of problems for which the model might be used. Reviewing model documentation and results is not sufficient peer review for many regulatory models. Because many stakeholders and others interested in the regulatory process do not have the capability or resources for a scientific peer review, they need to be able to have confidence in the evaluation process. This need requires a transparent peer review process and continued adherence to criteria provided in EPA's guidance on peer review. Documentation of all peer reviews, as well as evidence of the agency's consideration of comments in developing revisions, should be part of the model origin and history. (NRC, 2007, pp. 5–6)

elements from the evaluation framework because these activities are often not documented when a model is published (such as those that fall in the resources category). The two examples below illustrate a range of the elements in the framework.

Agent-Based Model of Potential Plug-in Hybrid Electric Vehicles (PHEVs) Market Adoption

Eppstein et al. (2011) describes an ABM of the potential market adoption of plug-in hybrid electric vehicles (PHEVs) that features spatial, social, and media influences. The model's purpose is to inform manufacturers and policy makers on the prioritization of investments toward potential leverage points and to identify combinations of policies that may be the most effective for PHEV market penetration.[4] In developing the ABM, however, Eppstein et al. recognized the need for additional data to inform the model. Thus, to strengthen the model, the developers conducted an extensive survey to gather and integrate data to the ABM (Krupa et al., 2014). As the researchers reworked the model with the new data, they generated results that could provide better insights for policy makers and manufacturers into the factors influencing the potential for PHEV market penetration (Eppstein et al., under review). Below is a detailed description of the process the researchers used in developing their ABM.

The original model by Eppstein et al. (2011) included agents who are individual vehicle consumers restricted to certain attributes.[5] When agents make decisions to buy a car in the model, they compare the relative costs and benefits of all pairs of vehicles and fuel types and then choose the most desirable vehicle. While agents think about their decisions, they may be susceptible to media and social influences. To put these agent attributes and decision rules into the model and determine their cross-correlations, Eppstein et al. (2011) drew on available data as well as on social science theories[6] and relevant literature.

The original model generated several findings relevant to PHEV

[4]Specifically, Eppstein et al. (2011) examined the effects of the following: gas prices; the ability of agents to consider fuel costs, PHEV purchase price, and rebates; PHEV all-electric battery range; consumer values regarding financial versus nonfinancial concerns in vehicle purchase; agent comfort threshold with the PHEV technology; social and media influence on PHEV market penetration; and fuel efficiency of the resulting fleet after 25 years.

[5]Agent attributes included age, annual salary, residential location, typical years of car ownership, annual vehicle miles traveled, vehicle age, fuel type, and fuel economy of current vehicle.

[6]The theories they used included threshold effects (Granovetter, 1978), homophily (McPherson et al., 2001), and conformity (Axelrod, 1997).

adoption.[7] Eppstein et al. (2011) also discussed a number of limitations of the ABM, including the lack of data for accurate parameterization and model realism. The modelers used data where possible to initialize agent attributes and simulations; where they did not have data, they tried to make reasonable yet simplifying assumptions. For example, the developers made many assumptions on spatial and inter-attribute cross-correlations and distributions, such as estimating the mean and standard deviation of the threshold distribution for new PHEV technology consumer adoption. These assumptions may not have necessarily been realistic, and they could have significantly affected model outcomes. Eppstein and colleagues did not claim that their model provided accurate quantitative predictions, but they stated that the findings offered preliminary insights into the combinations of policies and procedures that may be most effective for PHEV market penetration.

In order to provide more accurate parameterization and model realism, Eppstein et al. collected relevant quantitative data by administering an extensive consumer survey (Krupa et al., 2014). Each survey respondent corresponded one-to-one with an agent in the model so that each agent's attitudes and attributes, such as demographic information and susceptibility to social and media influences, were based on a real person. In this way, Eppstein et al. could populate the model with realistic (instead of assumed) distributions and cross-correlations of agent attributes. The survey included questions on different aspects of potential PHEV adoption barriers and attempted to fill in the holes left from the original model. Based on the analyses of the survey questions (Krupa et al., 2014), Eppstein et al. inserted agent vehicle purchasing decision rules in the model (Eppstein et al., under review). Data from the surveys revealed that many of the cross-correlations and estimates used in the original model, such as the standard deviation of the threshold distribution, were not accurate. The model developers continued to use some assumptions (e.g., rules for social network updates) in the modified model, but nowthey were equipped with more data, which resulted in different implementation decisions. Consequently, the updated model generated results that differed slightly from those of the original model.[8]

---

[7] Some of the findings included if there are sufficient potential early adopters, readily accessible estimates of lifetime vehicle fuel costs could be important for promoting PHEV market penetration; increasing gas costs could help people choose PHEV over traditional vehicles; temporal incentive programs like tax credits are not likely to have lasting effects on long-term fuel efficiency unless manufacturers are able to lower sticker prices after the rebates are discontinued; and increasing PHEV battery range may be an important leverage point.

[8] The results of the modified model indicated, among other things, that consumer uneasiness with the new PHEV technology was the biggest barrier to potential PHEV market penetration; that manufacturers and policy makers may need to take more action to help consumers feel

This model illustrates several of the elements outlined in the evaluation framework. Although the committee does not have information about the human or infrastructure resources for this model, the authors did strive to develop a model with the intended users—i.e., policy makers—in mind. The developers grounded their assumptions in theories during the conceptual phase of development. Although, as discussed, the model still contains assumptions, Eppstein and colleagues quantified and communicated the uncertainties and limitations of the model, provided additional data to better ground the model after the initial iteration of the model was completed, and incrementally developed the model, taking a key step in providing better insight into factors influencing the potential for PHEV market penetration. Although the initial model design did not properly represent the agents' behaviors, the authors made needed adjustments to improve the model for its intended purpose (exemplifying the necessary feedbacks in the evaluation framework presented in this chapter). The authors were clear on how the results of the model could be interpreted by policy makers, and where more information was needed.

SimAmph

A group of Australian researchers developed an ABM to study how individual perceptions, peer influence, and subcultural settings shape the use of psychostimulants and related harms among young Australians (Dray et al., 2012; Moore et al., 2009; Perez et al., 2012). The team was composed of modelers as well as experts in epidemiology, anthropology, economics, and drug policy. Within an interdisciplinary team, the researchers focused on collective design and incremental development of the model to address the study question. The team developed an ABM called SimAmph that iteratively integrated ethno-epidemiological data (Moore et al., 2009).

The researchers conducted both ethnographic and epidemiological studies simultaneously in three research sites, led by the appropriate experts on the team.[9] When developing the model, the ethnographers and epidemi-

---

more at ease with the new technology, whether it is through advertisements or well-publicized incentives; that many consumers choose used cars instead of new cars, whereas PHEVs are not likely to become part of an extensive used-car market anytime soon; that consumers may not feel limited when PHEVs are offered as only compact cars; that increases in gasoline prices may lead to small effects on PHEV market penetration (a finding that was contrary to the results from the first model); and that governmental and manufacturer rebates may allow PHEVs to be more competitive, but because many consumers may not know of the rebates, the rebates need to be more available until the prices of PHEVs decrease.

[9]The findings of the ethno-epidemiological data drawn from participant observation and in-depth interviews, and two surveys have been published (Green and Moore, 2009; Jenkinson et al., 2014; Siokou and Moore, 2008; Siokou et al., 2010).

ologists advised on the input data and the conceptual underpinnings of the ABM based on the findings of their studies, and the modelers asked questions, reworked the model, and conducted partial verification at each stage in the process. In addition, the team used secondary sources from national drug surveys as well as other qualitative research on similar populations to complement the findings of the ethno-epidemiological research and to further develop the model.

From these various sources, the researchers found that the use of psychostimulants among young Australians occurred mostly in the context of weekend partying and poly-drug use at licensed and other leisure venues. The researchers also learned that many young Australians were influenced by social relationships and the settings in which drug use took place. Using these findings, the researchers developed a model that included agents (young people) with particular attributes (e.g., socio-demographic characteristics, peer relationships) in various social settings who are able to access different types of drugs, have a set of friends whom they can exchange information with, such as drug experiences, and use drugs variably, depending on time and circumstance. The researchers set up rules, specifically concerning peer influence and health experience, that were designed to capture the dynamic process of the agents' use of psychostimulants (see Moore et al., 2009, for more details about the model). Over many iterations of model development, the researchers produced an ABM that could run such policy scenarios as the impact of pill testing (Moore et al., 2009) and the use of drug detection dogs by police and the dissemination of mass media prevention campaign (Dray et al., 2012).

SimAmph provides a good example of several of the criteria laid out in the evaluation framework. Having an interdisciplinary team in place from the outset allowed the researchers to explore many angles of the research question. Although SimAmph is simple and has several limitations,[10] the researchers integrated (or considered) concepts and data from relevant disciplines to capture and adequately justify the conceptual basis and inputs of the model while acknowledging the model's shortcomings. The team faced tensions brought on by the existence of multiple epistemologies rooted in different disciplines, but with ongoing, open dialogue throughout the project, the team was able to produce a model that integrates triangulated data and that begins to encapsulate and promote discussions concerning the complexity of drug use and policy (Moore et al., 2009). This type of interaction, which is highlighted in the evaluation framework, can help build a

[10] For example, the simulation was in a closed system that simplified a more complex reality of transient movements among individuals in drug scenes. For a comprehensive list, refer to page 70 of Perez et al. (2012).

strong conceptual framework for a model and increase the likelihood that the model will meet its intended purpose.

Several of the authors of SimAmph are part of the Drug Policy Modelling Program (DPMP), which created a series of models, including four ABMs (SimARC, SimDrug, SimDrugPolicing, and SimHero), that were designed to examine the effects of drug policies.[11] DPMP is tasked with generating new research evidence, translating evidence for policy makers, and studying how policy is made with teams that span many disciplines.[12] These goals are incorporated in the ABMs they have created. The team consults with policy makers to improve their use of the models and research. Although the model documentation does not include information on the financial resources available to DPMP, it is evident that input from an array of disciplines was considered and that the researchers sought critical human and knowledge resources during the course of model development. Because of the policy focus of DPMP, the researchers work with policy makers to ensure that the model is suitable for their purposes, and they regularly assign a "knowledge broker" to translate model findings into policy language and communicate the limitations of the modeled scenarios as well as the predictive ability of the model to the policy makers (MacDonald, 2012).[13] Because DPMP aims to ensure that modelers understand the needs of the model they are developing and to make certain that the models are used properly by policy makers, communication and translation strategies are considered throughout model development.

---

[11] See http://dpmp.unsw.edu.au/resource/models.

[12] These disciplines include complex systems science, criminology, economics, epidemiology, integration and implementation sciences, law, medicine, political science, psychology, public health, public policy, sociology, and systems thinking.

[13] Personal communication, P. Perez, A. Ritter, and Institute of Medicine staff, April 15, 2014.

Chapter 4 Annex

EXAMPLE OF APPLYING GRAMMAR TO DESCRIBE AGENT-BASED MODELS

The following Table 4-2 illustrates how the descriptive grammar presented in Table 4-1 could be applied to existing policy-relevant models. The grammar is meant to be descriptive only—it is not an evaluation of a model, but rather a systematic way to describe ABMs early in the model development process. The use of the grammar will improve communication between the model development team and the policy makers and help ensure that they are all in agreement about the goals and intended uses of the model. The models listed in Table 4-2 are described more fully in Chapter 4 (PHEV Market Adoption and SimAmph) and Chapter 5 (SnapDragon).

TABLE 4-2 Application of Descriptive Grammar to Three Policy-Relevant ABMs

	Models		
	PHEV Market Adoption[a]	SimAmph[b]	SnapDragon[c]
Basic Description			
Purpose	To inform policies affecting plug-in hybrid vehicle market penetration.	To test policies that could influence drug use and experience among young Australians.	To study the effects of tobacco control policies in a single- or multiple-tobacco product environment.
Breadth	Moderately broad	Very broad	Moderately narrow
Abstraction	Moderately realistic	Moderately realistic	Moderately abstract
Agents			
Type	Single type	Single type	Single type
Definition	Agents are consumers who make decisions about which vehicles to purchase	Agents are Australian youth who make decisions about drug use based on psychological and health status and social interactions	Agents are generic persons who have opinions about tobacco products and also tobacco use behaviors

TABLE 4-2 Continued

	Models		
	PHEV Market Adoption[a]	SimAmph[b]	SnapDragon[c]
Data and Theories			
Data—rules	Yes	Yes	No
Data—characteristics	Yes	Yes	Yes
Data—validation	No–data for validation not available	Yes—validated with data from the 2004 National Drug Strategy Household Survey	Very simple validation using social network data[d]
Theories	Social threshold effects, social science theories (principles of homophily and conformity)	Broad set of social science theories; developed ethnographic framework, Stage of Social Engagement	Opinion dynamics
Context			
Physical space	Abstract	Abstract	None
Social space	Simple	Simple	Simple
Physical dynamics	Static	Static	None
Social dynamics	Static	Static	Static (at the time of committee review)
Outcomes			
Primary outcome	Fleet fuel efficiency resulting from agent vehicle purchase choices	Individual drug use and population prevalence of drug-related harm and of regular drug use	User or nonuser of tobacco products
Proximal/distal outcome	Distal	Proximal/Distal	Proximal/Distal
Policy			
Policy definition	Effects of purchase rebates[e]	Effects of mass media drug prevention campaigns; effects of using drug-sniffing dogs	Introduction of non-specific communications campaign; introduction of new products
Policy realism	Realistic	Realistic	Abstract
Policy tests	Yes (although not a primary goal of study)	Yes	No (at the time of committee review)

continued

TABLE 4-2 Continued

	Models		
	PHEV Market Adoption[a]	SimAmph[b]	SnapDragon[c]
	Communications		
Model sharing	Collated results in the form of peer-reviewed papers and presentations.	Collated results in the form of peer-reviewed papers and presentations. Model code and documentation are available on a website.[f]	Some preliminary results have been presented at professional meetings; other aspects of the modeling process and outcomes have been presented to FDA; manuscripts have been submitted for publication.

[a]Sources: Eppstein et al., 2011; Krupa et al., 2014.
[b]Sources: Dray et al., 2012; Moore et al., 2009; Perez et al., 2012.
[c]Sources: Moore et al., in press a,b.
[d]These data were collected as part of NIH/NCI grant 3R01CA157577-02S1 (Extending a School-Based Cohort to Improve Longitudinal Modeling), Thomas W. Valente, principal investigator. This data collection was a follow-up to the Social Network Study cohort in 2010 through 2012 (Valente et al., 2013). The data are not yet published.
[e]Eppstein et al. (2011) did not identify specific policies to test from the beginning, but rather used the model to find key leverage points—that is, specific model parameters that, if changed, affected PHEV technology adoption—and then identified examples of potential government influence on the model parameters, through the form of a targeted policy. In addition to purchase rebates, other potential policy examples include gasoline taxes, tax breaks or other manufacturer incentives to keep PHEV sticker prices low, and public service announcements to educate consumers, among others.
[f]A version of SimAmph (and relevant documentation) is available at: http://cormas.cirad.fr/en/applica/simAmph.htm.

EVALUATION QUESTIONS DERIVED FROM THE EVALUATION FRAMEWORK FOR POLICY RELEVANT AGENT-BASED MODELS

Based on the evaluation framework presented on page 92, this document contains sample questions for each of the categories outlined. Although many of these questions would be of interest to any modeling effort, some questions are specifically applicable for ABMs, and many are geared toward informing models specific to tobacco control policies. The questions are intended for modelers, subject-matter experts, funders, policy makers, and other relevant collaborative members involved with developing or using the model. Before modeling begins, it is suggested that these actors select a reasonable number (e.g., three to five) of high-priority evaluation categories from the framework, develop questions within each category (potentially adapting the sample questions below), and build a tailored evaluation plan. If done properly, during and after model development, the evaluators (including independent third-party evaluators) would be able to understand the purpose of the model and apply the evaluation plan. Thus, the framework and associated questions are not meant to be used as a checklist but rather as a general guide that may help in determining if the model has fulfilled its objective. These questions do not reflect an evaluation of an actual ABM; however, many of these questions were considered by the committee as they reviewed the ABM developed for FDA (see Chapter 5). The questions are drawn from existing sources (ASPE, 2012; Caro et al., 2012; CREM, 2009; Grimm et al., 2006; Gurcan et al., 2011; Hammond, 2015; Kopec et al., 2010; Kuntz et al., 2013; NRC, 1991, 2007, 2012; Rochester, 2014; Šalamon, 2011; Wagner et al., 2010; Weinstein et al., 2001, 2003) as well as from committee expertise.

1. **Resources**
   a. Financial
      i. Were the model development, data acquisition, and model dissemination and data sharing funded at a level commensurate with the scope of the model?
      ii. Did the model developers have the required financial resources to reach the needs of the end users of the model?
   b. Infrastructure
      i. What hardware resources did the model developers use?
      ii. What software resources did the model developers use?
   c. Human
      i. Did the modeling team use an interdisciplinary team or approach when building and testing the model?
      ii. How were subject-matter experts involved (or not involved) in the model development?
      iiii. To what extent were relevant staff and stakeholders available (e.g., funders, policy makers, and end users) when building the model, especially in the conceptual development phase?
   d. Knowledge
      i. Were the specific policy or regulatory goals of the modeling project clearly described before model development began?
      ii. To what extent did the modeling team use (or at least take into account) the relevant studies and principles and frameworks in the area, not just knowledge of their own approach? That is, would all or some aspects of another approach be better suited to address the policy question or goal?

iii. What types of decisions could the model support (e.g., strategic planning, compliance, enforcement)?
iv. What kinds of data are available to support the model (e.g., epidemiologic, behavioral, public health system, tobacco industry)?

2. Activities

Internal Model Development Activities

a. Conceptual
   i. Why was the modeling method chosen (versus other approaches)? Were there other modeling methods that could have been used instead of or in tandem with this method?
   ii. How did the particular theoretical framework enhance or weaken the validity of the model results?
   iii. Was the level of abstraction employed in the model well justified, and did it match up well with the specific policies being examined?
   iv. Was a rationale presented for the overall scope and timeline of the model?
   v. What are the definitions of the major model components (e.g., agent characteristics, agent rules, environment, initiation, cessation, addiction, relapse)?
   vi. Did the model developers use appropriate theories to inform agent characteristics and interactions?

b. Model Implementation
   i. Did the model developers make full use of existing, relevant datasets? When empirical data were lacking, how was this accounted for in the model?
   ii. How are the assumptions supported (e.g., empirical evidence)?
   iii. Are social networks important for the specific model application? If so, were the social network structures and processes too simple (or too complex) for the model?
   iv. What kind of heterogeneity was captured? Did the model capture too little or too much?
   v. What temporal and spatial scales were used in the model, and were they appropriate for the presumed behaviors of the policies and agents?
   vi. What algorithms or mathematical methods are used in the model and how were they derived?
   vii. Were various evolving environmental scenarios, not just the status quo and past trends, considered in the model? What features were held constant?

viii. Is the model unreasonably complicated? (Are there, for example, too many parameters that increase model uncertainty? Did the modeling team consider trade-offs between the need for the model to be an accurate representation of the system of interest and the need for it to be reproducible, transparent, and useful for the regulatory decision?)

c. Model Testing
   i. What kinds of analyses were performed to quantify uncertainty?
   ii. Was the model output compared to empirical outputs under some specified time frame to ensure that the model captures real-world dynamics?
   iii. What problems and interesting or surprising model behaviors were identified, and how did the modeling team handle them?
   iv. How do the results compare to the results of other models addressing similar policy questions or having similar purpose?
   v. Do the results conform to or conflict with other relevant evidence and face validity?
   vi. How appropriate are the verification, validation, and calibration techniques used in the model?

d. Policy Testing
   i. How were the specific policies or processes operationalized within the modeling framework?
   ii. Were policies examined in isolation, or were multiple policies modeled and allowed to interact?

External Activities

e. Communications
   i. Were relevant stakeholders included in all aspects of the model development, or just at the end?
   ii. How were initial results shared with the stakeholders?
   iii. Were appropriate data and communication platforms developed for the model?
   iv. Were the model processes and results communicated in a manner that allows for reproducibility?
   v. If proprietary issues and requirements limited the communication of modeling information, were the costs and benefits of those limitations assessed or articulated?

f. Peer Review
   i. At what stages of the model development did the modelers seek peer review? What did the peer review involve (e.g., reviewing the conceptual framework of the model, running the model several times, comparing the model's results with known test cases, reviewing the model code)? How did the modelers incorporate the feedback into the model? Is there documentation of this?

3. Outputs

a. Frameworks
   i. How does the model design documentation describe all of the important details of the model implementation and testing process?
   ii. Does the model documentation include a write-up of model uncertainties, an interpretation of results, and considerations for maintenance of the model?
   iii. Did the authors provide a conceptual framework and causal map (this would be developed during the conceptual phase of model development)?
   iv. Did the authors clearly discuss the model's strengths and weaknesses and implications for tobacco control policy?

b. Development/Software Versioning System
   i. How did the modeling team use a management system to enhance model development? ("Management systems" are needed when building a model that requires a complicated software program.) How was progress documented?
   ii. Did the authors publish the model code and empirical databases?

c. Model Results
   i. What kinds of results were generated (e.g., morbidity, mortality, prevalence, DALYs [disability-adjusted life years])?
   ii. To what extent can the model address short-term, intermediate, and long-term effects?
   iii. Did the authors provide for a systematic storage of model output data?

d. Policy Results
   i. How useful are the model results for informing or setting priorities of future policy or regulatory activity (e.g., identification of promising policies, policy leverage points, implementation strategies)?
   ii. Does the model fulfill its designated task (i.e., address the specified policy goal(s))?

iii. How are the policy results translated and interpreted? Can the various audiences understand the model results, strengths, and limitations?

e. Communications
   i. Were the relevant stakeholders included in dissemination activities?
   ii. What kinds of multimedia platforms were used for dissemination?
   iii. Did contract restrictions or proprietary concerns inhibit dissemination?
   iv. Was a dissemination plan discussed with the funders?
   v. Were the model details and results clearly described? How accessible is the model?

4. Outcomes

a. Short-term
   i. How were preliminary results fed back into subsequent model improvements?
   ii. Based on model results, did policies and regulatory options change?
   iii. Who is going to use the model? How will it be applied? Do the end users have the expertise needed for using the model, or will they always need to partner with a contractor to use it?
   iv. How can this type of model be used to inform other models—for example, aggregate (compartment) models?
   v. How flexible is the model (i.e., capacity for the model to be modified or revised and applied to situations as new data arise or alternative objectives are specified)? What factors might trigger the need for major revisions, or what circumstances might prompt users to seek an alternative model?

b. Medium-term
   i. What was the return on investment for the modeling efforts? Are the results justified, given the amount of money invested and the amount of time taken to develop, test, and disseminate the model?
   ii. How has the sponsor (e.g., FDA) used the model's results to inform its own regulatory activities? Did the results help shape new regulatory and funding announcements?

iii. How relevant are the modeling results to the tobacco control field? Have the results informed tobacco control knowledge and influenced decisions among funders, regulators, policy makers, scientists?
iv. Have public health scientists collected new data to inform future model development and policy research?

c. Long-term
i. How has the sponsor or other stakeholders used the model to implement evidence-based tobacco control policies and regulation?
ii. How has the sponsor or other stakeholders used the model to improve population health via reducing product harms and addictiveness, preventing youth initiation, or increasing adult cessation?
iii. Did the model inform new promising avenues of research, study, or exploration?

REFERENCES

ASPE (Office of the Assistant Secretary for Planning and Evaluation). 2012. *Demystifying microsimulation meeting report*. Washington, DC: Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services.

Auchincloss, A. H., and A. V. Diez Roux. 2008. A new tool for epidemiology: The usefulness of dynamic-agent models in understanding place effects on health. *American Journal of Epidemiology* 168(1):1–8.

Axelrod, R. 1997. The dissemination of culture: A model with local convergence and global polarization. *Journal of Conflict Resolution* 41(2):203–226.

Badham, J. 2010. A compendium of modelling techniques. *Integration Insights #12*. http://i2s.anu.edu.au (accessed May 5, 2014).

Bammer, G. 2012. *Disciplining interdisciplinarity: Integration and implementation sciences for researching complex real-world problems*. Canberra, Australia: Australian National University Press.

Bammer, G., A. Michaux, and A. Sanson. 2010. *Bridging the "know-do" gap: Knowledge brokering to improve child wellbeing*. Canberra, Australia: Australian National University Press.

Bayarri, M. J., J. O. Berger, R. Paulo, J. Sacks, J. A. Cafeo, J. Cavendish, C.-H. Lin, and J. Tu. 2007. A framework for validation of computer models. *Technometrics* 49(2):138–154.

Berk, R. 2008. How you can tell if the simulations in computational criminology are any good. *Journal of Experimental Criminology* 4(3):289–308.

Bielak, A., A. Campbell, S. Pope, K. Schaefer, and L. Shaxson. 2008. From science communication to knowledge brokering: The shift from "science push" to "policy pull." In *Communicating science in social contexts*, edited by D. Cheng, M. Claessens, T. Gascoigne, J. Metcalfe, B. Schiele and S. Shi. Netherlands: Springer. Pp. 201–226.

Caro, J. J., A. H. Briggs, U. Siebert, and K. M. Kuntz. 2012. Modeling good research practices—overview: A report of the ISPOR-SMDM Modeling Good Research Practices Task Force—1. *Value in Health* 15(6):796–803.

CDC (Centers for Disease Control and Prevention). 1999. Framework for program evaluation in public health. *Morbidity and Mortality Weekly Report: Recommendations and Reports* 48(RR11):1–40.

———. 2007. *Evaluation guide: Developing and using a logic model.* http://www.cdc.gov/dhdsp/programs/spha/evaluation_guides/logic_model.htm (accessed May 5, 2014)

———. 2011. *Introduction to program evaluation for public health programs: A self-study guide.* Atlanta, GA: Centers for Disease Control and Prevention.

CORE (Cornell Office for Research on Evaluation). 2009. *The evaluation facilitator's guide to systems evaluation protocol.* Ithaca, NY: Cornell Digital Print Services.

CREM (Council for Regulatory Environmental Modeling). 2009. *Guidance on the development, evaluation, and application of environmental models.* Washington, DC: Office of the Science Advisor, Council for Regulatory Environmental Modeling, U.S. Environmental Protection Agency.

Dobbins, M., K. DeCorby, and T. Twiddy. 2004. A knowledge transfer strategy for public health decision makers. *Worldviews on Evidence-Based Nursing* 1(2):120–128.

Dray, A., P. Perez, D. Moore, P. Dietze, G. Bammer, R. Jenkinson, C. Siokou, R. Green, S. L. Hudson, and L. Maher. 2012. Are drug detection dogs and mass-media campaigns likely to be effective policy responses to psychostimulant use and related harm? Results from an agent-based simulation model. *International Journal of Drug Policy* 23(2):148–153.

Eddy, D. M., W. Hollingworth, J. J. Caro, J. Tsevat, K. M. McDonald, and J. B. Wong. 2012. Model transparency and validation: A report of the ISPOR-SMDM Modeling Good Research Practices Task Force—7. *Value Health* 15(6):843–853.

Eppstein, M. J., D. K. Grover, J. S. Marshall, and D. M. Rizzo. 2011. An agent-based model to study market penetration of plug-in hybrid electric vehicles. *Energy Policy* 39(6):3789–3802.

Eppstein, M. J., D. M. Rizzo, B. H. Y. Lee, J. S. Krupa, and N. Manukyan. Under review. National survey respondents as agents in a model of plug-in hybrid electric vehicle adoption.

Granovetter, M. 1978. Threshold models of collective behavior. *American Journal of Sociology* 83(6):1420–1443.

Green, R., and D. Moore. 2009. "Kiddie drugs" and controlled pleasure: Recreational use of dexamphetamine in a social network of young Australians. *International Journal of Drug Policy* 20(5):402–408.

Grimm, V., U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S. K. Heinz, G. Huse, A. Huth, J. U. Jepsen, C. Jørgensen, W. M. Mooij, B. Müller, G. Pe'er, C. Piou, S. F. Railsback, A. M. Robbins, M. M. Robbins, E. Rossmanith, N. Rüger, E. Strand, S. Souissi, R. A. Stillamn, R. Vabø, U. Visser, and D. L. DeAngelis. 2006. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling* 198(1):115–126.

Gurcan, O., O. Dikenelli, and C. Bernon. 2011. Towards a generic testing framework for agent-based simulation models. In *Proceedings of the Federated Conferece on Computer Science and Information System,* edited by M. Ganzha, L. Maciaszek, and M. Paprzycki. Los Alamitos, CA: IEEE Computer Society Press. Pp. 635–642.

Hammond, R. A. 2015. Considerations and best practices in agent-based modeling to inform policy. Paper commissioned by the Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation (see Appendix A).

Helbing, D., and S. Balietti. 2011. *How to do agent-based simulations in the future: From modeling social mechanisms to emergent phenomena and interactive systems design.* Tech. Rep. 11-06-024, Santa Fe, NM: Santa Fe Institute.

Hovelynck, J., A. Dewulf, G. François, and T. Taillieu. 2010. Interdisciplinary knowledge integration through group model building: Recognizing dualities and triadizing the conversation. *Environmental Science & Policy* 13(7):582–591.

Jenkinson, R., D. Jolley, and P. Dietze. 2014. "Weekend on the town": Discrete sessions of drug use for a sample of young psychostimulant users. *Drug and Alcohol Review* 33(4):428–435.

Jordan, G. B. 2010. A theory-based logic model for innovation policy and evaluation. *Research Evaluation* 19(4):263–273.

Kopec, J. A., P. Finès, D. G. Manuel, D. L. Buckeridge, W. M. Flanagan, J. Oderkirk, M. Abrahamowicz, S. Harper, B. Sharif, A. Okhmatovskaia, E. C. Sayre, M. M. Rahman and M. C. Wolfson. 2010. Validation of population-based disease simulation models: A review of concepts and methods. *BMC Public Health* 10(1):710.

Krupa, J. S., D. M. Rizzo, M. J. Eppstein, D. Brad Lanute, D. E. Gaalema, K. Lakkaraju, and C. E. Warrender. 2014. Analysis of a consumer survey on plug-in hybrid electric vehicles. *Transportation Research Part A* 64:14–31.

Kuntz, K., F. Sainfort, M. Butler, B. Taylor, S. Kulasingam, S. Gregory, E. Mann, J. M. Anderson, and R. L. Kane. 2013. *Decision and simulation modeling in systematic reviews*. Rockville, MD: Agency for Healthcare Research and Quality.

Langer, E. M., A. L. Gifford, and K. Chan. 2011. Comparative logic modeling for policy analysis: The case of HIV testing policy change at the Department of Veterans Affairs. *Health Services Research* 46(5):1628–1645.

Lempert, R. 2002. Agent-based modeling as organizational and public policy simulators. *Proceedings of the National Academy of Sciences of the United States of America* 99(Suppl 3):7195–7196.

MacDonald, D. 2012. Assessing the influence on drug policy of a program of drug policy research: The Australian Drug Policy Modelling Program 2006–2011. Paper presented at sixth annual conference of the international society for the study of drug policy, University of Kent, Canterbury, United Kingdom.

Marietto, M., N. David, J. Sichman, and H. Coelho. 2003. A classification of paradigmatic models for agent-based social simulation. In *Multi-agent-based simulation III*. Vol. 2927, Lecture notes in computer science, edited by D. Hales, B. Edmonds, E. Norling, and J. Rouchier. Germany: Springer Berlin Heidelberg. Pp. 193–208.

McPherson, M., L. Smith-Lovin, and J. M. Cook. 2001. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27(1):415–444.

Meyer, M. 2010. The rise of the knowledge broker. *Science Communication* 32(1):118–127.

Moore, D., A. Dray, R. Green, S. L. Hudson, R. Jenkinson, C. Siokou, P. Perez, G. Bammer, L. Maher, and P. Dietze. 2009. Extending drug ethno-epidemiology using agent-based modelling. *Addiction* 104(12):1991–1997.

Moore, T. W., P. D. Finley, N. S. Brodsky, T. J. Brown, B. Apelberg, B. Ambrose, R. J. Glass. In press a. Modeling education and advertising with opinion dynamics. *The Journal of Artificial Societies and Social Simulation*.

Moore, T. W., P. D. Finley, B. J. Apelberg, B. Ambrose, N. S. Brodsky, T. J. Brown, C. Husten, R. J. Glass. In press b. An opinion-driven behavioral dynamics model for addictive behaviors. *European Physical Journal B*.

Nicolson, C. R., A. M. Starfield, G. P. Kofinas, and J. A. Kruse. 2002. Ten heuristics for interdisciplinary modeling projects. *Ecosystems* 5(4):376–384.

NRC (National Research Council). 1991. *Improving information for social policy decisions—the uses of microsimulation modeling: Volume 1, Review and recommendations*. Washington, DC: National Academy Press.

———. 2007. *Models in environmental regulatory decision making*. Washington, DC: The National Academies Press.

———. 2012. *Assessing the reliability of complex models: Mathematical and statistical foundations of verification, validation, and uncertainty quantification*. Washington, DC: The National Academies Press.

Perez, P., A. Dray, D. Moore, P. Dietze, G. Bammer, R. Jenkinson, C. Siokou, R. Green, S. L. Hudson, and L. Maher. 2012. SimAmph: An agent-based simulation model for exploring the use of psychostimulants and related harm amongst young Australians. *International Journal of Drug Policy* 23(1):62–71.

Roberts, M., L. B. Russell, A. D. Paltiel, M. Chambers, P. McEwan, and M. Krahn. 2012. Conceptualizing a model: A report of the ISPOR-SMDM modeling good research practices task force—2. *Value in Health* 15(6):804–811.

Rochester, C. G. 2014. Developing simulation models for assessing effects of tobacco products. Paper presented at 2014 Tobacco Products Scientific Advisory Committee Meeting, Rockville, MD.

Šalamon, T. 2011. *Design of agent-based models: Developing computer simulations for a better understanding of social processes*. Czech Republic: Tomáš Bruckner.

Siokou, C., and D. Moore. 2008. “This is not a rave!”: Changes in the commercialised Melbourne rave/dance party scene. *Youth Studies Australia* 27(3):50–57.

Siokou, C., D. Moore, and H. Lee. 2010. “Muzzas” and “old skool ravers”: Ethnicity, drugs and the changing face of Melbourne’s dance party/club scene. *Health Sociology Review* 19(2):192–204.

Valente, T. W., K. Fujimoto, J. B. Unger, D. W. Soto, and D. Meeker. 2013. Variations in network boundary and type: A study of adolescent peer influences. *Social Networks* 35(3):309–316.

Wagner, W., E. Fisher, and P. Pascual. 2010. Misunderstanding models in environmental and public health regulation. *NYU Law Environmental Law Journal* 18:293–356.

Weinstein, M. C., E. L. Toy, E. A. Sandberg, P. J. Neumann, J. S. Evans, K. M. Kuntz, J. D. Graham, and J. K. Hammitt. 2001. Modeling for health care and other policy decisions: Uses, roles, and validity. *Value in Health* 4(5):348–361.

Weinstein, M. C., B. O’Brien, J. Hornberger, J. Jackson, M. Johannesson, C. McCabe, and B. R. Luce. 2003. Principles of good practice for decision analytic modeling in healthcare evaluation: Report of the ISPOR task force on good research practices—modeling studies. *Value in Health* 6(1):9–17.

5

Review of the Social Network Analysis for Policy on Directed Graph Networks Model

The U.S. Food and Drug Administration (FDA) commissioned the development of an agent-based model (ABM) through an interagency agreement with Sandia National Laboratories (SNL), with model development beginning in May 2010.[1] A major component of the statement of task provided by the FDA to this committee was to review the model, identify its strengths and weaknesses, and make recommendations for its improvement. This chapter describes this model, entitled Social Network Analysis for Policy on Directed Graph Networks (SnapDragon), and, where appropriate, applies the evaluation framework for policy-relevant ABMs presented by the committee in Chapter 4 to the model. Some of the model evaluation criteria cannot be applied in this chapter because the activities either happened before the committee's review (including many of the inputs) or else have not yet taken place, as the model is still undergoing development (many of the outputs and outcomes).[2] This chapter provides an analysis of the model and discusses its usefulness for informing tobacco control policy.

[1] SNL is a federally funded research and development center, operated and managed by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation. SNL functions as a contractor for the U.S. Department of Energy's National Nuclear Security Administration. The lab supports federal, state, and local government agencies, companies, and organizations. For more information, see http://www.sandia.gov/about.

[2] The framework presented by the committee in Chapter 4 was developed to assist FDA in the future development of policy-relevant ABMs and to provide a framework for the committee to use for its assessment of SnapDragon. As such, the evaluation framework captures all stages of the model development process, but not all of these can be used to assess SnapDragon at this time.

The committee's review is based on the SnapDragon model as it existed in July 31, 2014. The modeling team has continued to develop and test the model since that point,[3] but the committee did not review any features the modeling team added after July 2014 and did not base its review on changes the development team plans to incorporate in the future.

BACKGROUND

The SnapDragon model was developed for use by FDA to examine the impact of smoking control policies on certain population smoking metrics, such as prevalence as well as initiation and cessation rates. FDA first directed the model development team to use the model to explore the potential effects of various public education campaigns on the prevalence of tobacco use to help inform its public education efforts. To date, the work on SnapDragon has focused primarily on studying the effects of multiple competing high- and low-opinion messages in a network, illustrated through the study of education campaigns (SNL, 2014a).[4] The early stages of conceptual model development began in May 2010. Between the initial conception and the review of the model by this committee, the model development team developed both a single- and multiple-product model (see details below in model description), identified data needs, conducted sensitivity analyses, and presented their model at various conferences.[5] The model development team continues to develop SnapDragon, as the model is still in exploratory stages. The SnapDragon model has not yet been published in a peer-reviewed journal, but two manuscripts on the model were undergoing peer review in two different journals during the writing of this report.[6] In addition to the draft manuscripts, the committee received more

[3] In July 2014 and January 2015, draft descriptions of the committee's technical understanding of the model were sent to SNL for technical review. In their January 2015 response, the developers noted several changes to the model that occurred after July 31, 2014, and identified additional changes they planned to make. However, the review by the committee is based on the model as it existed on July 31, 2014. These documents are available upon request from the project public access file: http://www8.nationalacademies.org/cp/ManageRequest.aspx?key=49612.

[4] E-mail communication between the Institute of Medicine (IOM) and FDA staff, June 10, 2014. Available upon request from the project public access file: http://www8.nationalacademies.org/cp/ManageRequest.aspx?key=49612.

[5] E-mail communication between the IOM and SNL staff, July 14, 2014; available upon request from the project public access file: http://www8.nationalacademies.org/cp/ManageRequest.aspx?key=49612.

[6] Both manuscripts were accepted for publication at the end of this study. However, the manuscripts reviewed by the committee are available in the project public access file: http://www8.nationalacademies.org/cp/ManageRequest.aspx?key=49612. One manuscript, Modeling Education and Advertising with Opinion Dynamics, is dated May 2013 and was revised November 2013 (Moore et al., in press a); the second manuscript is titled An Opinion-Driven

information on SnapDragon from in-person presentations by the development team during two open meetings in February and June 2014 (SNL, 2014a,b) as well as from written question-and-answer documents produced by the committee and the model development team (draft manuscripts and other supporting documents are available upon request from the project public access file).[7]

SnapDragon Model Description

Based on the materials the committee received from the SnapDragon model developers (as outlined above), the committee offers the following description of SnapDragon.[8] SnapDragon is an ABM designed to study the effect of tobacco control policies in a single- or multiple-tobacco-product environment. The model distinguishes individuals by any number of characteristics, and in particular, according to their use of a variety of tobacco products, allowing for the investigation of the simultaneous use of different products. Currently, the model classifies individuals as either "users" or "nonusers" for each tobacco product under consideration. The user status is determined by an underlying construct termed "opinion." An opinion is an aggregate concept that captures the overall positive or negative attitude of a person toward a tobacco product. It is represented as a continuous variable with values between 0 and 1, with 0 standing for the most negative attitude a person can have toward a tobacco product, and 1 the most positive. Each individual carries an opinion about each tobacco product under consideration.

Opinions about a product can vary over time, influenced by the opinions of other individuals. SnapDragon explicitly models the time trajectory of individuals' opinions as a result of their interactions with other individuals, and the modeling choice is based on theory stemming from the field of opinion dynamics. In the model, individuals are connected to others through predefined social networks.[9] Connected individuals can affect each other's opinions if such opinions do not differ by more than a specified

---

Behavioral Dynamics Model for Addictive Behaviors, dated November 2013 and revised February 2014 (Moore et al., in press b).

[7] The project public access file is available at: http://www8.nationalacademies.org/cp/ManageRequest.aspx?key=49612.

[8] In addition, the model developers reviewed this section (SnapDragon Model Description) for technical accuracy.

[9] Erdős–Rényi (ER) graphs "were selected as a neutral illustrative framework for the results presented to the IOM. SnapDragon currently includes multiple classes of graphs, including scale-free, forest fire, community structure graphs drawn from Lancichinetti et al. (2008), transitivity-based graphs as proposed in Jackson and Rogers (2005), dynamic graphs, and regular graphs such as rings and lattices" (e-mail communication between the IOM staff and SNL staff, August 1, 2014; available upon request from the project public access file).

tolerance range. As time progresses, individuals continuously adjust their opinions toward a weighted average of the opinions of the individuals who can influence them. The weights (or plasticity values) represent the importance given to the opinion of particular individuals.[10] These plasticity values are not necessarily reciprocal, meaning that each individual in a connected pair can assign a different weight to the opinion of the other.

In the model, opinions drive tobacco use behavior, which is defined as being either a user or a nonuser of each of the specific tobacco products under consideration. Nonusers whose opinions about a tobacco product increase beyond a certain level (termed the initiation threshold) become users of that tobacco product. Users of a particular product whose opinions about such product fall below a certain level (termed the cessation threshold) become nonusers of the product. The initiation threshold is assumed to be above the cessation threshold, and the difference between these two levels indicates the degree of addiction of an individual to that particular product. An individual whose opinion is above the initiation threshold is a user; an individual whose opinion is below the cessation threshold is a nonuser; an individual whose opinion falls between the initiation and cessation thresholds could be either an addicted user if he or she has previously crossed the initiation threshold or else a nonuser if he or she has not.

Thresholds and opinions are determined by multiple factors. These determining factors, or determinants, can be identified to isolate their effect on policy interventions. In particular, SnapDragon test runs have been conducted to examine the effects of two particular determinants—risk perception and risk affinity—for hypothesized tobacco control policy interventions. Risk affinity is a personal attribute that defines the tendency of an individual to engage in risky activities. Other things being equal, the greater an individual's risk affinity is, the lower the initiation and cessation thresholds of a product will be. Risk perception is a component of an individual's opinion that measures the degree to which the individual perceives the product as harmful. Other things being equal, the greater the risk perception, the lower the individual's opinion of the product.

The SnapDragon model allows for the investigation of multiple tobacco products in the market, either by considering an individual to use different products simultaneously or by allowing an individual to switch between products. The model handles multiple product use by considering opinions and thresholds for each product simultaneously, with each product use determined by its own dynamics. Switching between two products is

[10] "Weights can therefore represent the closeness of the relationship (e.g., a best friend can be more highly weighted than friend), as well as the effectiveness of a media campaign" (e-mail communication between the IOM staff and SNL staff, August 1, 2014; available upon request from the project public access file).

handled in the following way: An individual can switch from product *A* to product *B* if the user's opinion of product *A* is between the initiation and cessation threshold for such a product. In such a case the difference between the product *A* initiation threshold and the user's opinion of product *A* (a "regret" factor) is added to (reinforces) the individual's opinion of product *B*. If this reinforced opinion exceeds the initiation threshold for product *B*, the individual switches from product *A* to product *B*.[11]

Interventions that can potentially influence individuals' behavior, such as tobacco control efforts, are modeled as modifying either the opinions of individuals about a certain tobacco product or the opinion thresholds that delimit possible user status. For example, price increases can be modeled by raising the initiation and cessation thresholds for a product. In contrast, a public health education campaign can be represented by adding a fictitious individual to the model's network (a "media node") who has a fixed low opinion of the product. This media node can influence the opinion of a certain number of individuals within the social network, but it is not influenced by them. In this case, the media node influences the opinions of the individuals it reaches, lowering the individuals' opinions of the product and potentially triggering a behavioral change.

Other types of interventions can be modeled in a similar way. For example, a tobacco product's advertising campaign can be represented by adding to the model a media node with a positive opinion about the product, while promotional price discounts can be modeled by lowering the product's initiation and cessation opinion thresholds. "SnapDragon is designed to incorporate multiple interventions in a scenario in order to model interactions and to analyze complementary and conflicting effects. Interventions can precede one another sequentially or run in parallel."[12]

The modeling team uses 2014 data from Tom Valente's high school networks study[13] to evaluate the empirical validity of two assumptions within the SnapDragon model: (1) that opinion influences behavior; and (2) that people are more likely to be friends with others who share their opinions about smoking. With regard to the former, the modeling team asserts that

[11] "The 'regret' factor can also be scaled (up or down) to reflect product characteristics such as substitutability. For example, if Product *B* is a less suitable replacement for Product *A*, then the additional opinion boost should be scaled down" (e-mail communication between the IOM staff and SNL staff, August 1, 2014; available upon request from the project public access file).

[12] E-mail communication between the IOM staff and SNL staff, August 1, 2014; available upon request from the project public access file.

[13] These data were collected as part of National Institutes of Health (NIH)/National Cancer Institute (NCI) grant 3R01CA157577-02S1 (Extending a School-Based Cohort to Improve Longitudinal Modeling), Thomas W. Valente, principal investigator. This data collection was a follow-up to the Social Network Study cohort in 2010 through 2012 (Valente et al., 2013). The data are not yet published.

these preliminary data are "consistent with opinion-to-behavior mapping in SnapDragon" (SNL, 2014b). With regard to the latter, the team members assert that the data are "consistent with the influence-network hypothesis" of SnapDragon (SNL, 2014b). "The data and analyses are considered preliminary, and ongoing analyses will be compared with analyses of Add Health and other data sets."[14]

SNAPDRAGON MODEL EVALUATION

In the remainder of this chapter the committee offers its assessment of the SnapDragon model. The committee focuses on two major evaluation categories outlined in Chapter 4—model suitability and technical best practices—as these are the categories for which the committee had adequate information with which to conduct an analysis. These two categories encompass the "activities" in the logic model presented in Chapter 4, particularly the conceptual development of the model, the model's implementation, and model testing. Some of the evaluation categories (such as communication and translation) are not yet relevant, as SnapDragon has not yet reached the later phases of model development. Before the model evaluation is presented, the chapter offers an overview of opinion dynamics, as opinion dynamics is the conceptual framework that drives the implementation of SnapDragon. Following this overview, the committee assesses the suitability of the opinion dynamics approach, as implemented in SnapDragon, to inform to tobacco control policy. (For example, does the opinion dynamics approach, as used, have face validity? Does the model incorporate relevant results from the literature in tobacco control?) Finally, the committee evaluates the technical aspects—namely the platform, parameters, and data use—of the model. (For example, has opinion dynamics been empirically validated for use in this context? Does it have predictive validity outside the field of tobacco control? Does the SnapDragon implementation have empirical validity? Have the developers demonstrated that the model's results agree well with known trajectories of smoking patterns following real-world interventions?) This assessment is based on the committee's collective expertise and on its interpretation of the supporting literature.

SnapDragon Model: Conceptual Overview

As noted earlier, opinion dynamics provides the underlying conceptual framework for the SnapDragon model. By basing SnapDragon on opinion dynamics, the modelers are making the explicit assumption that the

[14] E-mail communication between the IOM staff and SNL staff, August 1, 2014; available upon request from the project public access file.

dynamics that govern smoking initiation, cessation and, in general, product use are mainly dependent on the users' opinions about tobacco products and, further, that those opinions are influenced mainly by the interaction among individuals. The following section provides a brief overview of the opinion dynamics modeling approach.

Background on Opinion Dynamics

The goal of modeling opinion dynamics is to determine the opinion states in a population and the transitional processes between such opinion states (Castellano et al., 2009). Therefore, a common aim of opinion dynamics models is to identify how the opinions of individual agents are influenced by the opinions of neighboring agents and how they all converge to consensus.[15] Conceptually, opinion dynamics stems from sociological and social psychology theories (Cartwright and Harary, 1956; Heider, 1946) and studies (Asch, 1956; French, 1956) that focus on collective behavior and social influence and suggest that individual attitudes and behaviors tend to conform to the majority of the belonging group. The mathematical basis of opinion dynamics is derived from Ising spin models in statistical physics (Galam and Moscovici, 1991; Galam et al., 1982). Given that physics methods are being applied to describe social phenomena, opinion dynamics is generally regarded as an area of sociophysics (Castellano et al., 2009; Galam, 2008).

Over the past 15 to 20 years, as sociophysicists have actively worked in opinion dynamics (Castellano et al., 2009), they have developed several different implementation approaches. Some examples of these implementation approaches are the voter model (Clifford and Sudbury, 1973; Holley and Liggett, 1975), the majority rule model (Galam, 2002; Krapivsky and Redner, 2003), the Snzajd model (Stauffer, 2002; Sznajd-Weron, 2005), the cultural dissemination model (Axelrod, 1997), and the bounded confidence model (Deffuant et al., 2000; Hegselmann and Krause, 2002). (For more information on all of these models, see Castellano et al., 2009, and Xia et al., 2011). These distinct types of opinion dynamics models can differ in the way that they represent opinions (e.g., continuous versus discrete), in their local rules of interaction (e.g., averaging of opinions), and in their underlying structure (e.g., regular lattice, dimensions, and networks). Using various implementations of opinion dynamics–based rules, sociophysicists have incorporated opinion dynamics into models across a number of domains. For instance, opinion dynamics has been applied to voter be-

[15] This is not always the case, as shown in some reports in the literature, such as the study by Yildiz et al. (2011) in which the aim is to model stubborn agents that never come to an agreement.

havior (Ben-Naim et al., 2003) and consensus building in politics (Galam, 2008), the diffusion of agricultural practices in Europe (Weisbuch et al., 2002), the spread of propaganda (Carletti et al., 2006), tribal and gendered leadership in Afghanistan (Moore et al., 2012; Schubert et al., 2011), extremist group dynamics and terrorism (Backus and Glass, 2006; Deffuant et al., 2002), and marketing strategies (Martins et al., 2009; Schulze, 2002; Sznajd-Weron et al., 2008). However, opinion dynamics models have not yet been applied to tobacco control or to the wider field of public health.

Opinion dynamics has brought new perspectives to the social sciences and has pointed to new questions and directions for research (Castellano et al., 2009; Galam, 2008; Lorenz, 2007; Xia et al., 2011); however, several opinion dynamics experts have noted that opinion dynamics has not been properly empirically validated and that most attempts to do so have only used election data (Moss, 2008; Sobkowicz, 2009; Weisbuch, 2007). Castellano et al. (2009) argued that the field needs to focus on the development of better defined quantitative models of consensus formation, which can describe this phenomenon in a more objective way, beyond addressing the mere qualitative question of when and how people agree/disagree. Furthermore, as Moussaïd et al. (2013, pp. 1–2) wrote, "it is difficult to track and measure how opinions change under experimental conditions, as these changes depend on many social and psychological factors such as the personality of the individuals, their confidence level, their credibility, their social status, or their persuasive power." Existing opinion dynamics models tend to start either from plausible criteria on the effect of social interactions on opinion changes or from established social theories, but there has been a minimal effort to compare the predictions of the models with data on real social dynamics. This makes it difficult to model opinion changes or to propose a meaningful validation of the basic mechanisms in opinion dynamics.

Opinion Dynamics and SnapDragon

The model developers use opinion dynamics as the conceptual foundation of SnapDragon, but have made some adjustments in its implementation. In traditional opinion dynamics models, opinion is a general term for beliefs (Carletti et al., 2006; Martins, 2008) or attitudes (Huet et al., 2008; Jager and Amblard, 2005) that are held by individuals. In SnapDragon, "opinion" represents an integrated value of positive and negative attitudes and perceptions of an individual toward a tobacco product.[16] Additionally,

[16] "'Opinion' is an integrated view of a product that is the result of multiple influences and perceptions. In our model it is an acquired behavioral disposition toward smoking which is based upon an aggregation of salient conceptual components and evaluations. Opinion is unidimensional and can range from 0 (lowest opinion of a product) to 1 (highest opinion of

the SnapDragon development team chose to incorporate certain technical elements from various opinion dynamics models. These features include a bounded confidence approach within a social network topology and also media and behavior components, as detailed below.

The SnapDragon development team applied the bounded confidence approach to the model. A widely known bounded confidence opinion dynamics model has been developed by Deffuant and Weisbuch (Deffuant et al., 2000; Weisbuch et al., 2002). In bounded confidence opinion dynamics models, the opinions of agents are represented as continuous variables, ranging between 0 and 1. As in many other types of opinion dynamics models, the opinions of agents in SnapDragon can, over time, be influenced by other agents in the environment, either through random connections in a well-mixed, non-networked population or else by interactions within a social network topology, with the latter being what the SnapDragon model uses. However, in bounded confidence models, agents interact with each other only when their opinions are close together—that is, within certain tolerance bounds; if their opinions are very different from one another, they do not interact (see Equation 1). In the final stationary state, one, two, or more clusters emerge, signifying consensus, polarization, or a fragmentation of opinions, respectively. Eventually, the opinions of all agents within a given cluster converge to the same value.

$$\left|x_i(t) - x_j(t)\right| \le \varepsilon_i$$

EQUATION 1 *i* represents an individual, and *j* represents a neighbor to *i*. ε_i is the opinion tolerance bound for individual *i*. The equation specifies the range of opinion to which individual *i* might be receptive to interact with a given neighbor (Moore et al., in press b).

The model developers apply these general concepts of bounded confidence to SnapDragon but alter specific elements from the Deffuant–Weisbuch model. Although the Deffuant–Weisbuch model uses a bounded confidence model of non-directed interactions in well-mixed populations, the modeling team implemented directionality imposed by a network topology, so that the interaction between two agents is not necessarily reciprocal. Within this network structure, opinion clusters are formed based on the tolerance values of various individuals (see Equation 1). As time goes by,

a product). While opinion represents an aggregation of factors, it is not a mathematical summation of measured quantities. It is a model parameter used to represent positive and negative affective and utilitarian components that might influence a person's view of using a tobacco product" (e-mail communication between the IOM staff and SNL staff, April 3, 2014; available upon request from the project public access file).

networks that contain agents with low tolerance (that is, who are less open to influence) will spur the formation of many small fragmented clusters, and networks that contain agents with high tolerance (more open to influence) will move toward a large cluster consensus value.

The SnapDragon developers also deviated from the Deffuant–Weisbuch model in their use of an averaging function, which determines how an agent updates its opinions when interacting with other agents. Specifically, instead of applying a pairwise averaging function that captures randomized discrete interactions, the modelers implemented a rule that calls for agents to average the opinions (weighted by the plasticity values) of all the neighboring agents that satisfy the bounded confidence condition, a technique used in Hegselmann and Krause's bounded confidence opinion dynamics model (2002). In other words, with the model's averaging rule, agents move their opinions toward the weighted (by the plasticity values) average opinion of all agents that lie within their tolerance range (see Equation 2).

$$x_i(t+1) = x_i(t) + \frac{1}{|S_i|}\sum_{j \in S_i} \mu_{ij}[x_j(t) - x_i(t)]$$

EQUATION 2 Where x_i and x_j are as described in Equation 1, $x_i(t)$ is i's current opinion, and $x_j(t)$ is the current opinion of neighbor j. $x_i(t+1)$ is the opinion value of individual i at the next time step. When applied to a directed social network, S_i consists of the out-degree neighbors of individual i with cardinality. μ_{ij} is the plasticity value associated with the relationship between individual i and neighbor j (Moore et al., in press a,b).

In addition to using bounded confidence, the modeling team also incorporated media nodes into SnapDragon. In particular, they relied on the work of Carletti and colleagues (2006), who extended the Deffuant–Weisbuch model to model the effects of propaganda, in which the media act to target opinions and influence tolerance levels. The SnapDragon modeling team adopted this idea of media influence, but again, as described above, instead of assuming a well-mixed population in which the media interact with all individuals in the population at the same time, they defined those interactions within the constraints imposed by a social network. Thus, the media are integrated into the social network topology. A media node may have an effect on an influential member of a social network who, subsequently, will have an impact on other members of that social network. However, while the media nodes in SnapDragon have the ability to influence agents, they are not themselves subject to influence.

Another major component of SnapDragon is the connection between opinions and behavior. The model developers reference the Continuous Opinions and Discrete Actions (CODA) model through an update rule used by Martins (2008). The CODA model assumes that agents update their opinions by observing the actions of surrounding neighbors. The SnapDragon model adopts this notion of linking opinions to behavior, but instead of the agents observing the behaviors of others and subsequently changing their opinions, agents' opinions (which are susceptible to the influence of other agents) drive behavior change. Therefore, the modeling team integrates a concept reminiscent of CODA into SnapDragon, but it makes significant changes conceptually and does not apply the same implementation strategies.

Model Suitability of SnapDragon

In this section, the committee reviews the suitability of SnapDragon for its intended use (see the Chapter 4 evaluation framework for details regarding model suitability). Given this model, what sorts of policies and outcomes are amenable to modeling by SnapDragon? Although SnapDragon has been designed to evaluate a wide range of tobacco products, for ease of exposition the committee comments on how the structure of the model can accommodate known facts about smoking behavior. Models that describe smoking behavior have traditionally classified individuals by various demographic characteristics (e.g., age and gender) and smoking characteristics (the widely used tobacco control models to date are compartmental/aggregate models).[17] Usually, individuals in these models are categorized as never-smokers, current-smokers, or former-smokers, further classifying smokers by the number of cigarettes smoked per day and former smokers by years-quit (HHS, 2014; Jeon et al., 2012; Levy et al., 2006; Mendez et al., 1998). These classifications are important because smoking-associated health risks are known to vary by age, gender, smoking status, smoking intensity, and, in the case of former smokers, by years-quit.

Many of the existing tobacco control models follow groups of individuals through time. Up to a certain age, individuals have a certain chance of starting to smoke. As time progresses, a smoker has the opportunity to quit or to continue smoking, while a former smoker has a certain chance of relapsing. In most of these existing models, the rates to start or quit smoking are exogenously supplied. SnapDragon characterizes individuals as being either users or nonusers of tobacco products. Therefore, the model in its current form can track prevalence of product use, but it cannot accurately determine health risks, because often a great proportion of tobacco-related morbidity and mortality falls on former users of combustible products (HHS, 2014). Although determining health risks was not listed as one of

[17] For a review of many existing tobacco control models, see the 2014 Surgeon General's report, Appendix 15.1 (HHS, 2014).

the purposes of SnapDragon, if the Center for Tobacco Products (CTP) plans to use SnapDragon as a stand-alone model, this will be a limitation. If CTP plans to use SnapDragon only to inform population models, it will not be a limitation of the model.

SnapDragon attempts to model the process of initiation and cessation as being driven by social interactions. Instead of inputting initiation and cessation rates that have been determined outside the model, the model tries to derive these figures endogenously, using a hypothesis of how these processes are generated. That is, SnapDragon attempts to explain the dynamics (i.e., how the system changes over time) inherent in tobacco use as a result of a convergence of opinions about specific tobacco products through the interaction among individuals in the population, guided by the opinion dynamics formulation discussed in the previous section. This overarching assumption supports the use of an agent-based framework to implement the model, as individual interactions are unique to the social network structure in which they occur.

Postulating a simple mechanism at the individual level to explain the multiple emergent complexities of tobacco use observed at the macro level is elegant and appealing, but several elements in SnapDragon's formulation either do not conform to existing knowledge or defy face validity.

First, the model does not consider a feedback mechanism from behavior to opinion. It is almost certain that the experience of using a particular tobacco product would influence the user's opinion about such product. For example, when individuals first start smoking, their prior opinions about cigarettes are likely to be altered by the particular experiences of the product. About one-third to one-half of all adolescents in the United States have ever smoked part or all of a cigarette (HHS, 2012), but a substantial proportion of those adolescents who ever smoked do not progress to regular smoking (ALA, 2010; CDC, 1998). It is conceivable that a portion of these youths only tried cigarettes for experimentation, without any intention of continuing to use the product, but it is more than likely that a significant number of the youths who tried cigarettes and did not progress to regular smoking were deterred by their personal experiences with the product. It is also known that adolescents who experiment with menthol cigarettes are more likely to become regular smokers than those who start smoking regular cigarettes (Nonnemaker et al., 2013), indicating that a specific feature of the product influences subsequent behavior. Similarly, it is known that cessation rates increase after age 50, when smokers start experiencing the adverse effects of their behavior (Mendez et al., 1998). All these examples point toward product features and use influencing subsequent behavior independently of social pressures. As SnapDragon takes into account only the modification of behaviors by opinions and not vice versa, it seems likely that the model is missing an important feedback mechanism—that is, from behavior to opinion.

Second, it is very unlikely that opinions about tobacco products are transmitted independently of individuals' behavior toward such products, as SnapDragon stipulates. This formulation could lead to highly unrealistic scenarios. For example, the model implies that two individuals with the same opinion about a tobacco product could exhibit different behaviors (user and nonuser) because of the addiction factor or personal differences in initiation and cessation thresholds. However, these two individuals with different behaviors will exert the same influence on the agents with whom they connect because they will transmit the same opinion.

It is certainly conceivable that imitation of smoking behavior could play a role in tobacco use adoption. In fact, the Bass model (1969), a well-established marketing model of the diffusion of goods in the market, proposes that the rate of adoption of a new product is determined by a set of independent self-initiator individuals, followed by a "contagion" or imitation process that depends on the volume of the product already in the market. In SnapDragon, however, the imitation component happens indirectly, by individuals sharing their opinions about a product, rather than their behavior. Because opinions are not influenced by behaviors in the model, a growth or decline in the number of tobacco users in the population will not affect the rates at which new adopters are generated. Similarly, observed quitting behavior cannot be imitated directly in the model. For example, if individuals in a group are near their cessation threshold and a slight decrease in their opinion levels (triggered perhaps by a policy) makes them quit simultaneously, their observed behavior would not produce an additional effect on other individuals beyond the initial adjustment of attitudes triggered by the policy.

The committee has not found any references in the health field literature that support the dynamics suggested by the opinion dynamics formulation, which are the underpinnings of SnapDragon (that is, the way SnapDragon describes how opinions evolve over time and how these opinions trigger actions). Diffusion models have been proposed in the marketing literature to explain the dynamics following the introduction of new products in the market (Mahajan et al., 1990), but these models have relied on imitating the adoption of the product rather than the diffusion of the underlying attitude toward such products, which may or may not trigger the adoption behavior.

Third, the rationale behind the modeling choice of making interacting opinions converge to a weighted average is not clear. This is clearly a modeling choice by the developers of SnapDragon, as opinion dynamics offers a number of ways by which opinions of different interacting agents can get reconciled (Castellano et al., 2009; Xia et al., 2011), including as one possibility the convergence to an average. This modeling choice, when applied to smoking behavior, can lead to inconsistencies between the model's results and observed facts. For example, studies suggest that there is almost no smoking initiation after age 26 (for example, among 30- to 34-year-olds, 89.8 percent of smokers initiated by age 19, and 99.2 percent by age 26 [HHS, 2012; IOM, 2015]). This indicates, following the logic implied by SnapDragon, that nonsmokers' opinion of tobacco smoking never rises above their initiation threshold after age 26, regardless of the potential multiple interactions with positive opinions about tobacco use throughout their lifetimes. This implies either that nonsmokers have a substantially higher initiation threshold than smokers or that the assumption of potentially converging opinions about tobacco through social interactions is not likely to be accurate. If it were, we would observe smoking initiation (albeit small) at all ages, due to the individuals' multiple encounters with positive messages about tobacco use throughout their lives. It is likely that other mechanisms, not reflected in SnapDragon, play an important role in modifying smoking behavior as people age (such as those identified in Chapter 2).

Fourth, another aspect of the model that defies face validity is the lack of a credible behavior for former smokers. The model does not consider relapses, and it is difficult to envision how it would. For individuals to quit, their opinions will have to run below the cessation threshold. It is known that many quitters relapse after a period of time because they continue to crave nicotine. The SnapDragon formulation would imply that former smokers' opinions about smoking would have to increase beyond their initiation level after quitting, triggered by interactions with other agents, which is a very unrealistic scenario.

Finally, while models based on opinion dynamics have been able to replicate the equilibrium patterns of a number of socially driven processes (Clifford and Sudbury, 1973; Holley and Liggett, 1975), the committee has not found applications in which the specific time path to equilibrium has been empirically validated. Estimation of time paths is important in tobacco control because the evaluation of policies usually involves the determination of discounted benefits and costs wrought by specific interventions. As the dynamics of smoking behavior carry much inertia, the full effects of tobacco control interventions may take a long time to realize, and thus the time trajectory of smoking rates becomes very relevant. As currently designed, the model is not suitable for addressing long-term dynamics but only the short-term impacts of policy interventions. Even for short-term assessments of policies, it is doubtful that SnapDragon will be able to generate more than qualitative scenarios at best, given that a realistic parameterization of the model is not likely to be feasible. For example, it would be very challenging to estimate individuals' baseline opinions and action thresholds, because smokers have different levels of addiction that may depend on genetics, smoking intensity, and age, among other factors (HHS, 2010).

Use of Data in SnapDragon

SnapDragon is meant to represent key aspects of the real-world process of tobacco use, particularly initiation and cessation. Data could thus serve as inputs to the model, ensuring that agents are realistic representations of persons. Data could also confirm whether the model is able to replicate or predict real-world patterns of initiation and cessation among individuals or populations. Data are critical at many if not all stages of model development. As discussed below, the current SnapDragon model does not use much data. Although SnapDragon is still in the early stages of development and testing, and the modeling team has outlined some areas where they plan to collect or use existing data (Moore et al., in press b; SNL, 2014a),[18] data could have played a more central role in informing the model during its early stages of development.[19] At least three types of data could be used to inform SnapDragon or future plans for its parametrization or testing: stylized facts that offer qualitative benchmarks, individual-level data on personal characteristics, and quantitative aggregated data. (Additional data needs for an ABM are discussed in Chapter 6.)

The most basic type of data that could be used in a model is stylized facts[20] that offer qualitative benchmarks. The ability to replicate stylized facts is a minimum bar that any model should be able to clear. Although such replication means that a model is able to capture general features of the real world, the qualitative nature of such facts precludes precision of the type that would convince policy makers of a model's validity. The SnapDragon modeling team mentions a number of known facts about tobacco use that could be used as stylized facts to inform the model (Moore et al., in press a,b), but SnapDragon incorporates only a small range of relevant and salient stylized facts (for example, peer influence in smoking initiation) to inform and validate the model. This is important because it affects the data used to inform the conceptual underpinnings of the model. The modeling team has used opinion dynamics to inform the model, but stylized facts such as varied individual quitting processes at different ages and changing peer influence by age are not included. The model has a great

[18] The modeling team noted that the initial stages of model development focused on model structure and that later development will incorporate more realistic data (SNL, 2014a).

[19] See also communication between the IOM and SNL staff, June 25, 2014; available upon request from the project public access file.

[20] Stylized facts are structural observations or a "set of properties, common across many instruments, markets and time periods . . . observed by independent studies" (Cont, 2001, p. 223).

deal of flexibility, but the relevant stylized facts are not used to inform the base model. Although these data could be incorporated in SnapDragon at later stages of development, it would have been more informative to do so early in model development.

The second type of data that could inform SnapDragon is individual-level data on the distribution of attributes of agents in the population, including the health behaviors of interest (tobacco use) as well as demographic information (age, gender, race, socioeconomic status) and other relevant agent attributes. Individual-level data would include the multiple characteristics of individual agents, which are likely to be correlated to one another. Data might be aggregated (for example, if the joint distribution of agent characteristics were known). Such data are readily available in multiple sources commonly used in health behavior research, such as the National Health and Nutrition Examination Survey[21] and the Behavioral Risk Factor Surveillance System (BRFSS).[22] Individual level data are commonly used in ABMs to create a one-to-one correspondence between agents and real-world persons (North and Macal, 2007). These data may also be used to monitor individual trajectories, which could serve as ground truth[23] against which to compare modeled trajectories. SnapDragon currently does not use individual-level data to specify agent characteristics.

A third type of data is quantitative contextual or aggregated data. Such data might arise from the aggregation of nationally representative surveys of the individuals just described. Data may permit an examination by geographic context, such as with the Tobacco Use Supplement of the Current Population Survey[24] or state-level data available in the BRFSS. The social network context of tobacco use is also available in some datasets (such as from Add Health), although such data are harder to come by (as described in greater detail below). As above, such data could be used as an input to initialize the model or as ground truth for model validation.[25] State-level comparisons between model outputs and real-world trends would increase confidence in the models' ability to capture the real-world data-generating process (Windrum et al., 2007). At present, SnapDragon makes limited use of data aggregated at the national level to calibrate initiation thresholds and addiction factors in order to produce smoking prevalence around observed levels. Such data are not used by SnapDragon to model differences seen across geographic regions or social networks. Although the commit-

[21] See http://www.cdc.gov/nchs/nhanes.htm (accessed March 2, 2015).

[22] See http://www.cdc.gov/brfss (accessed March 2, 2015).

[23] "Ground truth" refers to any data that capture the empirical process under investigation.

[24] See http://appliedresearch.cancer.gov/tus-cps (accessed March 2, 2015).

[25] Validation is defined as "the process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model" (AIAA, 1998, p. 3).

tee acknowledges that the SnapDragon developers are not yet at the stage of incorporating such data into the model (SNL, 2014a),[26] the committee cannot fully assess whether the model is behaving correctly because it does not reproduce observed facts.

SnapDragon is currently implemented at the level of a hypothetical population with 250 agents. One common goal of ABMs is the simulation of higher-level behavior that is emergent, such as the clustering of smoking behavior within networks that arises from peer influence. SnapDragon builds up from individual agents' behaviors and thresholds to look at patterns of smoking behavior (e.g., the prevalence of smoking) at the aggregate level of a high school in the United States or a small network in a larger community consisting of friends and friends of friends. This may present challenges for validation, especially if ground-truth data are not available at the level of interest. For example, the model may be calibrated to individual-level data of stages of tobacco use initiation, but it may be the case that only population-level data on tobacco consumption are available. Aggregated individual-level data could be compared by county or state to see if the patterns match (Berk, 2008). The Valente data[27] include 20 answers to opinion-related questions, some of which are related to attitudes toward tobacco use, but the data are cross-sectional[28] and closely tied to smoking behavior itself, and they are not currently incorporated into the model. The Valente data, though confirming that there is a clustering of these attitudes in the network, merely confirm a stylized fact. The SnapDragon development team identified the need for more longitudinal data; however, the type of data needed to inform SnapDragon is generally not available.

The uses of data are most extensively described by the SnapDragon modeling team in two places: on Table 1 of the draft journal manuscript (Moore et al., in press b) and in a presentation to the committee (slide 33, SNL, 2014b). The model developers also discussed some assumptions in a response to committee questions.[29] These descriptions show that the initial model is thin on data inputs and validation against external sources.

[26]See also communication between the IOM and SNL staff, January 26, 2015; available upon request from the project public access file.

[27]These data were collected as part of NIH/NCI grant 3R01CA157577-02S1 (Extending a School-Based Cohort to Improve Longitudinal Modeling), Thomas W. Valente, principal investigator. This data collection was a follow-up to the Social Network Study cohort in 2010 through 2012 (Valente et al., 2013). The data are not yet published.

[28]The data collected by Valente are longitudinal, but the 20 additional questions added on as part of NIH/NCI grant #CA157577-02S1 were collected only during the final year of data collection.

[29]E-mail communication between the IOM staff and SNL staff, January 21, 2014; available upon request from the project public access file.

Parameters are chosen to demonstrate model dynamics, rather than allowing a single parameter to dominate the model's behavior. In addition to the parameters specified in the draft manuscripts describing the SnapDragon model (Moore et al., in press a,b), the team has added two more parameters: (1) "risk affinity" to make the agents more heterogeneous and (2) "risk perception" to allow for product switching (SNL, 2014b). These parameters also are not based on data.

Calibration, Verification, and Validation

To ensure that a model is valid and that it accurately represents the real world for its intended use, a model must go through calibration,[30] verification,[31] and validation processes at various points in model development. At this point in time, SnapDragon is very general and flexible. What the team described as verification entails a comparative analysis with empirical research and sensitivity analysis to determine what drives the model's behavior (SNL, 2014a).[32] These exercises are limited to internal validation or calibration of model parameters to replicate real-world results—what Berk (2008, p. 291) calls "internal quantitative credibility," rather than confirmation that the model's data-generating process is the same as the real-world data-generating process (Windrum et al., 2007). The SnapDragon modeling team members report that they have plans to conduct "parameter analysis and uncertainty quantification to make sure the parameters are consistent with knowledge of the system."[33] However, it is not clear whether even these exercises would constitute Berk's external quantitative credibility—that is, comparisons of model output with real-world test data not used to develop and calibrate the model. The evaluations conducted so far have been "nearly data-free" (Berk, 2008, p. 293) relations between model output and ground truth that use very few stylized facts. Furthermore, these stylized facts are only qualitative rather than, say, quantitative values that draw from actual trends in smoking initiation and cessation over time. Missing so far is the search for areas where the model

[30] Calibration is "the process of adjusting numerical or physical modeling parameters in the computational model for the purpose of improving agreement with experimental data" (AIAA, 1998, p. 13).

[31] Verification is "the process of determining that a model implementation accurately represents the developer's conceptual description of the model and the solution to the model" (AIAA, 1998, p. 3). This includes code verification (Does the code correctly implement the intended algorithms?) and solution verification (accuracy in which the algorithms solve the mathematical-model equations for the specified quantity of interest) (NRC, 2012).

[32] See also communication between the IOM and SNL staff, June 25, 2014; available upon request from the project public access file.

[33] E-mail communication between the IOM staff and SNL Staff, January 21, 2014, page 3; available upon request from the project public access file.

does well and where it fails in making predictions and then using these analyses to refine and improve the model.

Specific Issues That Arise in Modeling Networks

In the examples that the modeling team shared with the committee, SnapDragon employed an Erdős–Rényi random graph; the modeling team reports that it plans to employ other stylized networks in future work. Other models have used real-world networks, such as airline routes (Epstein et al., 2007) and traffic patterns (Eubank et al., 2004), which are especially relevant for airborne infectious diseases. The use of real-world networks has the advantage of replacing assumptions about network structure with the actual network. It would be relatively straightforward for SnapDragon to use real networks as an input (e.g., Add Health or the data collected by Valente[34]), which would have the advantage of including agent attributes within the social network context. The disadvantage of inputting a fixed network is that the network-generative process and dynamics are not captured, which may be important if both peer selection and influence processes operate, as has been suggested for smoking behavior (Schaefer et al., 2012). Network dynamics are increasingly being incorporated in models, for example, to model behavioral changes in response to an epidemic outbreak (Epstein et al., 2008; Meloni et al., 2011) and network-behavior coevolution in smoking (Schaefer et al., 2013).

IMPLICATIONS FOR THE SNAPDRAGON MODEL

Summary of Findings and Conclusions

The SnapDragon model presents a novel framework for dealing with the complexities of tobacco use behavior. The developers of SnapDragon, which uses opinion dynamics methods, have suggested that it could be applied for a number of tobacco control policy applications, but the underlying assumptions of the model (as discussed in this chapter) suggest that this is unlikely. The committee statement of task calls for recommendations for improvement of SnapDragon, if needed, and although some changes could be made to address some of the weaknesses identified in this report, doing so would lead to the creation of a new model. SnapDragon does not encompass essential facts from the tobacco research literature, and many of

[34]These data were collected as part of NIH/NCI grant 3R01CA157577-02S1 (Extending a School-Based Cohort to Improve Longitudinal Modeling), Thomas W. Valente, principal investigator. This data collection was a follow-up to the Social Network Study cohort in 2010 through 2012 (Valente et al., 2013). The data are not yet published.

its assumptions lack face validity. In addition, the required data to inform the parameters in SnapDragon have not yet been identified, and the model has not yet reached the stage of model validation for broad application to tobacco control policy.

It is true that SnapDragon has the necessary flexibility to reproduce certain observed facts about tobacco use behavior by manipulating plasticity values and action thresholds. However, this could be problematic because, as Laine (2006, p. 37) writes,

> Overly flexible models, for instance ones with many free parameters, can be easily made to fit all these anomalies, byproducts of errors and noise, without capturing the regularities underlying the behavior. A model like this does not really inform us about the interesting patterns that may exist in the population, but just reflects the idiosyncrasy present in each individual sample. This is called overfitting.

SnapDragon is a very flexible model, but it currently lacks sufficient modeling structure to be informative. Therefore, the committee has not included recommendations for improvement. Key findings and conclusions regarding SnapDragon are below:

Conclusion 5-1: As SnapDragon presumes that opinions may modify behavior but that behavior does not modify opinion, the committee concludes that the model is missing an important feedback mechanism from behavior to opinion.

Finding 5-1: The committee finds that the representation of behavior in SnapDragon does not align with what is currently known about tobacco use and dependence.

Conclusion 5-2: The committee concludes that the modeling decision of making interacting opinions about tobacco converge to a weighted average is not supported by evidence and is unlikely to be an accurate representation of tobacco use behavior.

Finding 5-2: Whereas some other models based on opinion dynamics have been able to replicate the equilibrium patterns of socially driven processes, the committee has not found applications in which the specific time path to equilibrium has been empirically validated.

Finding 5-3: The committee finds that there has been no assessment of SnapDragon's ability to accurately predict initiation, prevalence, or cessation.

Conclusion 5-3: The committee concludes that a realistic parameterization of SnapDragon would be hard to achieve, so it is unlikely that the model will be able to generate credible assessments of policies.

Recommendation 5-1: SnapDragon should not be pursued by the Center for Tobacco Products as an aid for regulatory decision making.

Chapters 2, 3, and 4 offer findings, conclusions, and recommendations to assist CTP in the development of ABMs in the future. Chapter 6 offers guidance on inputs and implementation for ABM at CTP, drawing on lessons learned from the review of SnapDragon and other models.

REFERENCES

AIAA (American Institute of Aeronautics and Astronautics). 1998. Guide for the verification and validation of computational fluid dynamics simulations (AIAA g-077-1998(2002)). Reston, VA: American Institute of Aeronautics and Astronautics, Inc.

ALA (American Lung Association). 2010. *Children and teens*. http://www.lung.org/stop-smoking/about-smoking/facts-figures/children-teens-and-tobacco.html (accessed March 4, 2015).

Asch, S. E. 1956. Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs: General and Applied* 70(9):1–70.

Axelrod, R. 1997. The dissemination of culture: A model with local convergence and global polarization. *Journal of Conflict Resolution* 41(2):203–226.

Backus, G. A., and R. J. Glass. 2006. An agent-based model component to a framework for the analysis of terrorist-group dynamics. Albuquerque, NM: Sandia National Laboratories.

Bass, F. M. 1969. A new product growth for model consumer durables. *Management Science* 15(5):215–227.

Ben-Naim, E., P. L. Krapivsky, F. Vazquez, and S. Redner. 2003. Unity and discord in opinion dynamics. *Physica A: Statistical Mechanics and Its Applications* 330(1–2):99–106.

Berk, R. 2008. How you can tell if the simulations in computational criminology are any good. *Journal of Experimental Criminology* 4(3):289–308.

Carletti, T., D. Fanelli, S. Grolli, and A. Guarino. 2006. How to make an efficient propaganda. *Europhysics Letters* 74(2):222.

Cartwright, D., and F. Harary. 1956. Structural balance: A generalization of Heider's theory. *Psychological Review* 63(5):277–293.

Castellano, C., S. Fortunato, and V. Loreto. 2009. Statistical physics of social dynamics. *Reviews of Modern Physics* 81(2):591–646.

CDC (Centers for Disease Control and Prevention). 1998. Selected cigarette smoking initiation and quitting behaviors among high school students—United States, 1997. *Morbidity and Mortality Weekly Report* 47(19):386–389

Clifford, P., and A. Sudbury. 1973. A model for spatial conflict. *Biometrika* 60(3):581-588.

Cont, R. 2001. Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance* 1(2):223–236

Deffuant, G., D. Neau, F. Amblard, and G. Weisbuch. 2000. Mixing beliefs among interacting agents. *Advances in Complex Systems* 3(1–4):87–98.

Deffuant, G., F. Amblard, G. Weisbuch, and T. Faure. 2002. How can extremism prevail?: A study based on the relative agreement interaction model. *Journal of Artificial Societies and Social Simulation* 5(4):1.

Epstein, J. M., D. M. Goedecke, F. Yu, R. J. Morris, D. K. Wagener, and G. V. Bobashev. 2007. Controlling pandemic flu: The value of international air travel restrictions. *PLoS ONE* 2(5):e401.

Epstein, J. M., J. Parker, D. Cummings, and R. A. Hammond. 2008. Coupled contagion dynamics of fear and disease: Mathematical and computational explorations. *PLoS ONE* 3(12):e3955.

Eubank, S., H. Guclu, V. A. Kumar, M. V. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang. 2004. Modelling disease outbreaks in realistic urban social networks. *Nature* 429(6988):180–184.

French, J. R. P. 1956. A formal theory of social power. *Psychological Review* 63(3):181–194.

Galam, S. 2002. Minority opinion spreading in random geometry. *The European Physical Journal B* 25(4):403–406.

———. 2008. Sociophysics: A review of Galam models. *International Journal of Modern Physics C* 19(3):409–440.

Galam, S., and S. Moscovici. 1991. Towards a theory of collective phenomena: Consensus and attitude changes in groups. *European Journal of Social Psychology* 21(1):49–74.

Galam, S., Y. Gefen, and Y. Shapir. 1982. Sociophysics: A new approach of sociological collective behaviour. I. Mean-behaviour description of a strike. *Journal of Mathematical Sociology* 9(1):1–13.

Hegselmann, R., and U. Krause. 2002. Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation* 5(3)1–33.

Heider, F. 1946. Attitudes and cognitive organization. *The Journal of Psychology* 21(1): 107–112.

HHS (U.S. Department of Health and Human Services). 2010. *How tobacco smoke causes disease: The biology and behavioral basis for smoking-attributable disease: A report of the Surgeon General.* Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention, Office on Smoking and Health.

———. 2012. *Preventing tobacco use among youth and young adults: A report of the Surgeon General* Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

———. 2014. *The health consequences of smoking—50 years of progress: A report of the Surgeon General.* Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

Holley, R. A., and T. M. Liggett. 1975. Ergodic theorems for weakly interacting infinite systems and the voter model. *The Annals of Probability* 3(4):643–663.

Huet, S., G. Deffuant, and W. Jager. 2008. A rejection mechanism in 2D bounded confidence provides more conformity. *Advances in Complex Systems* 11(4):529–549.

IOM (Institute of Medicine). 2015. *Public health implications of raising the minimum age of legal access to tobacco products*. Washington, DC: The National Academies Press.

Jackson, M. O., and B. W. Rogers. 2005. Search in the formation of large networks: How random are socially generated networks? Working paper. Pasadena, CA: California Institute of Technology, Division of the Humanities and Social Sciences.

Jager, W., and F. Amblard. 2005. Uniformity, bipolarization and pluriformity captured as generic stylized behavior with an agent-based simulation model of attitude change. *Computational & Mathematical Organization Theory* 10(4):295–303.

Jeon, J., R. Meza, M. Krapcho, L. D. Clarke, J. Byrne, and D. T. Levy. 2012. Chapter 5: Actual and counterfactual smoking prevalence rates in the U.S. population via microsimulation. *Risk Analysis* 32:S51–S68.

Krapivsky, P. L., and S. Redner. 2003. Dynamics of majority rule in two-state interacting spin systems. *Physical Review Letters* 90(23):238701.

Laine, T. 2006. Agent-based model selection framework for complex adaptive systems. https://scholarworks.iu.edu/dspace/bitstream/handle/2022/7247/umi-indiana-1469.pdf?sequence=1 (accessed March 2, 2015).

Lancichinetti, A., S. Fortunato, and F. Radicchi. 2008. Benchmark graphs for testing community detection algorithms. *Physical Review E* 78(4):046110.

Levy, D. T., J. E. Bauer, and H. R. Lee. 2006. Simulation modeling and tobacco control: Creating more robust public health policies. *American Journal of Public Health* 96(3):494–498.

Lorenz, J. A. N. 2007. Continuous opinion dynamics under bounded confidence: A survey. *International Journal of Modern Physics C* 18(12):1819–1838.

Mahajan, V., E. Muller, and F. M. Bass. 1990. New product diffusion models in marketing: A review and directions for research. *The Journal of Marketing* 54(1):1–26.

Martins, A. C. 2008. Continuous opinions and discrete actions in opinion dynamics problems. *International Journal of Modern Physics C* 19(4):617–624.

Martins, A. C., C. de B. Pereira, and R. Vicente. 2009. An opinion dynamics model for the diffusion of innovations. *Physica A* 388(15):3225–3232.

Meloni, S., N. Perra, A. Arenas, S. Gómez, Y. Moreno, and A. Vespignani. 2011. Modeling human mobility responses to the large-scale spreading of infectious diseases. *Scientific Reports* 1:62.

Mendez, D., K. E. Warner, and P. N. Courant. 1998. Has smoking cessation ceased?: Expected trends in the prevalence of smoking in the United States. *American Journal of Epidemiology* 148(3):249–258.

Moore, T. W., P. D. Finley, R. J. Hammer, and R. J. Glass. 2012. Opinion dynamics in gendered social networks: An examination of female engagement teams in Afghanistan. In *Social computing, behavioral-cultural modeling and prediction*, edited by S. J. Yang, A. M. Greenberg, and M. Endsley, Vol. 7227. Germany: Springer Berlin Heidelberg. Pp. 69–77.

Moore, T. W., P. D. Finley, N. S. Brodsky, T. J. Brown, B. Apelberg, B. Ambrose, R. J. Glass. In press a. Modeling education and advertising with opinion dynamics. *The Journal of Artificial Societies and Social Simulation*.

Moore, T. W., P. D. Finley, B. J. Apelberg, B. Ambrose, N. S. Brodsky, T. J. Brown, C. Husten, R.J. Glass. In press b. An opinion-driven behavioral dynamics model for addictive behaviors. *European Physical Journal B*.

Moss, S. 2008. Alternative approaches to the empirical validation of agent-based models. *Journal of Artificial Societies and Social Simulation* 11(1):5.

Moussaïd, M., J. E. Kämmer, P. P. Analytis, and H. Neth. 2013. Social influence and the collective dynamics of opinion formation. *PLoS ONE* 8(11):e78433.

Nonnemaker, J., J. Hersey, G. Homsi, A. Busey, J. Allen, and D. Vallone. 2013. Initiation with menthol cigarettes and youth smoking uptake. *Addiction* 108(1):171–178.

North, M. J., and C. M. Macal. 2007. *Managing business complexity : Discovering strategic solutions with agent-based modeling and simulation*. Oxford, UK: Oxford University Press.

NRC (National Research Council). 2012. *Assessing the reliability of complex models: Mathematical and statistical foundations of verification, validation, and uncertainty quantification*. Washington, DC: The National Academies Press.

Schaefer, D. R., S. A. Haas, and N. J. Bishop. 2012. A dynamic model of US adolescents' smoking and friendship networks. *American Journal of Public Health* 102(6):e12–e18.

Schaefer, D. R., J. Adams, and S. A. Haas. 2013. Social networks and smoking: Exploring the effects of peer influence and smoker popularity through simulations. *Health Education and Behavior* 40(1 Suppl):24S–32S.

Schubert, J. L., T. W. Moore, P. D. Finley, R. Hammer, and R. J. Glass. 2011. Modeling tribal leadership dynamics: An opinion dynamics model of Pashtun leadership selection. Albuquerque, NM: Sandia National Laboratories. http://www.sandia.gov/CasosEngineering/docs/CSSS_Schubert_Tribal_Dynamics. pdf (accessed March 2, 2015).

Schulze, C. 2002. Advertising effects in Sznajd marketing model. *International Journal of Modern Physics C* 14(1): 95–98. http://arxiv.org/pdf/cond-mat/0207181.pdf (accessed March 2, 2015).

SNL (Sandia National Laboratories). 2014a. Presentation to the Committee on Assessment of Agent-Based Models to Inform Tobacco Product Regulation. Paper read at first meeting of the Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation, Washington, DC.

———. 2014b. Using SnapDragon to examine a two-tobacco-product system: The potential impact of risk perception on tobacco use prevalence. Paper read at third meeting of the Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation, Washington, DC.

Sobkowicz, P. 2009. Modelling opinion formation with physics tools: Call for closer link with reality. *Journal of Artificial Societies & Social Simulation* 12(1):11.

Stauffer, D. 2002. Sociophysics: The Sznajd model and its applications. *Computer Physics Communications* 146(1):93–98.

Sznajd-Weron, K. 2005. Sznajd model and its applications. *Acta Physica Polonica B* 36(8): 2537–2547.

Sznajd-Weron, K., R. Weron, and M. Włoszczowska. 2008. Outflow dynamics in modeling oligopoly markets: The case of the mobile telecommunications market in Poland. *Journal of Statistical Mechanics: Theory and Experiment* (11):P11018.

Valente, T. W., K. Fujimoto, J. B. Unger, D. W. Soto, and D. Meeker. 2013. Variations in network boundary and type: A study of adolescent peer influences. *Social Networks* 35(3):309–316.

Weisbuch, G. 2007. Social opinion dynamics. In *Econophysics and sociophysics*, edited by B. K. Chakrabarti, A. Chakraborti, and A. Chatterjee. Germany: Wiley & Sons. Pp. 339–366.

Weisbuch, G., G. Deffuant, F. Amblard, and J. P. Nadal. 2002. Meet, discuss, and segregate! *Complexity* 7(3):55–63.

Windrum, P., G. Fagiolo, and A. Moneta. 2007. Empirical validation of agent-based models: Alternatives and prospects. *Journal of Artificial Societies and Social Simulation* 10(2):8.

Xia, H., H. Wang, and Z. Xuan. 2011. Opinion dynamics: A multidisciplinary review and perspective on future research. *International Journal of Knowledge and Systems Science* 2(4):72–91.

Yildiz, E., D. Acemoglu, A. Ozdaglar, A. Saberi, and A. Scaglione. 2011. Discrete opinion dynamics with stubborn agents. *ACM Transactions on Economics and Computation* 1(4):19.

6

Data and Implementation Needs for Computational Modeling for Tobacco Control

The committee noted in the previous chapter that many types of data can be used to inform tobacco research and modeling. Three uses of data were highlighted: qualitative data and "stylized facts" that offer qualitative benchmarks; individual-level data on personal characteristics; and quantitative aggregated data. Empirical data, such as the results from cross-sectional studies, can be useful for indicating patterns of tobacco use in specific settings. Other types of model inputs, such as theoretical models and grounded theories that conceptualize social patterns and structures, qualitative data, and heuristics are also important to consider. In this chapter, the committee provides a high-level overview of existing tobacco-use data sources and identifies data gaps. The committee then discusses inputs and data sources for agent-based models (ABMs) that build off of the discussion of data use in the SnapDragon model (Moore et al., in press a,b) in Chapter 5, including types of data sources, types of network data, and future data collection needs. Different types of agents that could be included in individual-level models, from molecules and cells to individuals and institutions, are also discussed. The chapter finishes with recommendations for the future implementation of computational models at the U.S. Food and Drug Administration's (FDA's) Center for Tobacco Products (CTP).

EXISTING TOBACCO USE DATA SOURCES

Data are often critical at many, if not all, stages of model development. This section outlines existing data sources and identifies data gaps for tobacco control that could be filled to better understand the evolving tobacco landscape and to inform tobacco control models.

National, state, and local surveillance and evaluation systems primarily collect data on tobacco use behaviors and may also gather information on knowledge and attitudes about pro-tobacco and anti-tobacco influences, effects of tobacco use, and other important risk factors and health outcomes (CDC, 2014). These surveillance resources often vary in their timing, sampling methods, data collection modes, participation rates, and operational definitions and questions regarding tobacco use, initiation, and cessation. For example, the National Youth Tobacco Survey developed by the U.S. Centers for Disease Control and Prevention (CDC) uses self-administered surveys in classrooms to collect nationally representative data biennially on middle and high school youth's tobacco-related beliefs, attitudes, behaviors, social norms, and exposure to pro- and anti-tobacco influences. Every 3 years the Tobacco Use Supplement to the Current Population Survey (TUS–CPS) uses household interviews and telephone follow-ups to capture both national and state data on the age of initiation, secondhand smoke exposure, attitudes toward smoke-free policies, and cessation behavior among young adults and adults. (Box 6-1 provides an overview of national and state survey tools that include information on tobacco use; for more detailed information see CDC, 2014.) Although these surveys rely on different methodologies and are cross-sectional, they have provided general evidence on tobacco use and have offered insight for use in the planning, implementation, and evaluation of tobacco control programs as well as in policy making over the past few decades.

The National Longitudinal Study of Adolescent Health (Add Health) collects data from a nationally representative sample of U.S. adolescents who were in grades 7–12 during the 1994–1995 school year and includes, among many other topics, survey questions on tobacco use (Harris et al., 2009).[1] In 1994, Add Health collected nationally representative behavioral and network data on a baseline "core" sample of more than 90,000 students, including an "in-home" subsample drawn from the core who received more extensive interviews (n = 12,105); of these in-home respondents, 3,702 attended 1 of 16 "saturation schools" where a near-complete social network could be mapped out using answers to the questionnaires (Harris, 2013). Add Health also includes data on family, neighborhood, community, and schools; it is one of the few sources to provide data on both social networks and tobacco use.

[1] Only one cohort of adolescents was selected and followed into adulthood. The study collected follow-up data in 1994–1995, 1996, 2001–2002, and 2007–2008 using in-home interviews. Both public-use and restricted-use datasets are available. See http://www.cpc.unc.edu/projects/addhealth/data for more information (accessed March 2, 2015).

BOX 6-1
Tobacco Data Sources and Tools

National and State Surveys and Tools
- Adult Tobacco Survey
- Alaska Native Adult Tobacco Survey
- American Indian Adult Tobacco Survey
- Behavioral Risk Factor Surveillance System
- Hispanic/Latino Adult Tobacco Survey Guide
- Monitoring the Future
- National Adult Tobacco Survey
- National Health and Nutrition Examination Survey
- National Health Interview Survey
- National Survey on Drug Use and Health
- National Youth Tobacco Survey
- Pregnancy Risk Assessment Monitoring System
- School Health Policies and Practices Study
- School Health Profiles
- Smoking-Attributable Morbidity, Mortality, and Economic Costs
- State Tobacco Activities Tracking and Evaluation System
- Tobacco Use Supplement to the Current Population Survey
- Youth Risk Behavior Surveillance System
- Youth Tobacco Survey

Registries and Vital Statistics
- National Program of Cancer Registries
- National Vital Statistics System

Health Systems and Clinical Settings
- Healthcare Effectiveness Data and Information Set
- Hospital discharge data
- National Ambulatory Medical Care Survey
- National Hospital Ambulatory Medical Care Survey
- National Mental Health Services Survey
- National Quitline Data Warehouse
- National Survey of Substance Abuse Treatment Services

Sales Data
- Information Resources, Inc.
- Scanner data
- Substance Abuse and Mental Health Services Administration
- Tax revenue data
- U.S. Food and Drug Administration compliance checks

National, State, and Local Policy Tracking
- ACTIVE Life Tobacco Free Worksite Survey
- American College Health Association College Campus Tobacco Cessation and Prevention Survey

continued

BOX 6-1 Continued

American Lung Association's State Legislated Actions on Tobacco Issues
American Nonsmokers' Rights Foundation: U.S. Tobacco Control Laws Database
California Student Tobacco Survey
California Tobacco Use Prevention Education Evaluation Teacher Survey
CDC School Health Profiles
Worksite and Restaurant Smoking Policy Questionnaires and Guide

Media Tracking

Adobe SiteCatalyst
Arbitron
Cision
Clicktracks Optimizer
DataSift
Facebook Insights
Gnip
Google Analytics
HootSuite
Legacy Media Tracking Survey and Legacy Media Tracking Online
LexisNexis
Nielsen
Pinterest
Radian6
Sysomos
Topsy
Webalyzer
YouTube Analytics

Global Survey Tools

Global Adult Tobacco Survey
Global Health Professions Student Survey
Global School Personnel Survey
Global School-Based Student Health Survey
Global Youth Tobacco Survey

Tobacco Industry Monitoring

Network of the National Cancer Institute
New Product Watch, funded by Tobacco Surveillance, Epidemiology, and Evaluation
Project SMART Money of California State Department of Public Health
Retail Advertising Tobacco Survey
University of California at San Francisco Tobacco Control Archives

SOURCE: Adapted from CDC, 2014, which contains details on each of these sources.

The smoking and social network data available in Add Health have been used in numerous studies. For example, Pollard and colleagues found that membership in "higher-use" trajectories of tobacco smoking, as these adolescents moved into adulthood, were predicted by the number of perceived best friends who smoked and by changes in the numbers of these friends (Pollard et al., 2010). Several analyses of Add Health have employed the stochastic actor-based model SIENA developed by Snijders and colleagues (2010), which simultaneously models the social network change process and the peer influence process. Schaefer and colleagues (2012) found that students in a single Add Health saturation school smoked more frequently if their peers smoked, and they were also more likely to choose peers who smoked if they themselves smoked. Using the Add Health saturation schools, Lakon and colleagues (2014) considered the effects of parental influences as well and found that smoking by parents and peers increased the probability of an adolescent's smoking. The SIENA model is amenable to simulation and could serve as a basis for computational experiments for smoking prevention, as was done by Schaefer and colleagues (Haas and Schaefer, 2013, 2014).[2] Other studies using network and smoking data include the six-country European Smoking Prevention Framework studies by Mercken and colleagues (2009), a study of online social networks supporting tobacco cessation (Cobb et al., 2010), and studies by Valente and colleagues (2006, 2013).

Data from the Population Assessment of Tobacco and Health (PATH) Study could provide useful data in the future. PATH is a national-cohort longitudinal study of tobacco use and how it affects health in the United States. Sponsored by the National Institutes of Health and FDA, PATH began in 2011 and is a prospective study that will follow an estimated 46,000 U.S. household residents age 12 years and older (PATH, 2015a). The study's goals include explaining various aspects of tobacco use patterns and characterizing the natural history of tobacco dependence, cessation, and relapse. However, the PATH study is still in the early phases, and data from it are not yet available.[3] PATH will collect some data that are not routinely collected in other data sources. For example, PATH will identify trends in tobacco use patterns, including the use of new products, dual use,

[2] It is important to note some of the limitations of the SIENA model. For example, the model is limited to network change and peer influences on behavior, which is just one component of the model. Furthermore, SIENA is designed to fit a simulation model to data, and thus results (parameter estimates for network or behavior change) may not be generalizable to out-of-data scenarios. Finally, the model requires an initial network configuration.

[3] "The field test for the PATH Study took place between November 2012 and February 2013. Baseline data collection, which will last for 15 months, began in September 2013; the second annual data collection begins mid-October 2014 and will be followed by at least one additional data-collection wave" (PATH, 2015b).

poly use and switching; it will monitor changes in risk perceptions and other attitudes, such as social acceptability and individual preferences; and it will assess differences among and within critical subgroups, including youth, young adults, daily users, racial/ethnic minority groups, and users of new tobacco products, among others.

In addition to national and state surveys, tobacco-related information can be gathered from a variety of other sources, even if many of these sources are dedicated to other topics. Cancer registries, vital statistics, and medical records[4] offer data on health status and outcomes, such as incidence data on smoking-related morbidity and mortality. Quitline data warehouses collect information on the use and success of quitlines and identify knowledge gaps in order to inform the design of new strategies that can improve cessation services. Mass media and social media trackers can gather data on the level of influence of both anti- and pro-tobacco advertisements and campaigns as well as on tobacco-related beliefs, attitudes, social norms, and behaviors, particularly among youth (CDC, 2014). Finally, consumer purchase data have been collected and analyzed to assess trends in purchasing in order to identify patterns relevant to specific geographic locations and demographics characteristics of consumers and also to assess the impact of specific marketing strategies. One source of these data are the Nielsen data (2014), which can be purchased to understand better how and where specific products are selling[5] (see, for example, Amerson et al., 2014; NYSDOH, 2011; Terry-McElrath et al., 2011).

As the tobacco landscape has evolved in recent years, the need for different types of data has grown. After the enactment of the Tobacco Control Act—and in response to emerging trends in tobacco use—FDA and CDC began including detailed questions on nonconventional tobacco products in the 2012 National Youth Tobacco Survey (Apelberg et al., 2014). However, most surveys still focus on cigarettes, and the data sources are not available for every state (CDC, 2014). The surveys that have included questions on other tobacco products still lack the quality, depth, and breadth to capture data on the effects of multiple product use, substitution, and branding on initiation, cessation, addiction, and tobacco-related disparities among population groups (Delnevo, 2014; Mermelstein, 2014). There also continues to be a gap in the data on the interacting effects of multiple tobacco control

[4]Some tobacco studies have used medical records as data sources. For example, to study the relationship between passing smokefree indoor policies and incidence of myocardial infarction, Hurt and colleagues (2012) used medical records from the Mayo Clinic. Potentially, larger health care datasets could be used for future tobacco research.

[5]There are restrictions on how those data can be disseminated, but government agencies now purchase these data to assess local and national sales trends because of their relevance to a variety of outcomes (e.g., increases in sales of tobacco in a location might be linked with increased health care costs in that same area) (Amerson et al., 2014).

policies (Farrelly, 2009). Finally, network data for tobacco use—that is, the salient social connections between potential or current users and their peers, family members, and others who may influence tobacco use—are almost completely lacking, including data on special populations such as minority groups and high-risk groups such as those with mental illness. Such data could provide a better understanding of the influence of social networks and social context on tobacco use and on the behavior change process involved. Given the changing tobacco landscape, it is likely there will be an increasing need for detailed yet timely and accurate data for informing tobacco control efforts nationwide. Data needs for ABMs are discussed in more detail later in this chapter.

DATA NEEDS FOR FUTURE MODELING EFFORTS

Although various types of existing data sources related to tobacco use can be used to inform and strengthen ABMs, these sources do not contain all of the relevant agent attributes, behaviors, and social and spatial interactions related to tobacco use. As noted above, Add Health data are commonly used to study peer influences on smoking behavior. However, the Add Health baseline data lack detailed information on the mechanisms underlying peer selection. Also, its tobacco-related questions, which are concerned only with smoking and chewing/snuff, capture limited information (CPC, 1998). Furthermore, the biological and clinical data collected by Add Health are not comprehensive, especially in the first two waves of the study. The lack of data on networks and smoking can make modeling the social interactions that influence tobacco use a challenge.

Other existing data sources could also be used to inform computational models but pose some challenges as well (North and Macal, 2007, p. 240). The tobacco industry has collected much data on the uptake of smoking and the effectiveness of marketing (see Cummings et al., 2002, as well as the Legacy Tobacco Documents Library[6] for historical industry documents), which could be used to inform computational models. Sifting through these documents and finding those most relevant to inform computational models used to guide regulatory efforts could be difficult, however (Bero, 2003; Cruz, 2009). Another approach would be to try to maximize the use of available administrative data from states and regions, but, except in unusual circumstances, this information is not likely to contain many of the behaviors and interactions wanted. Alternatively, one could combine

[6]The Legacy Tobacco Documents Library is a digital archive of tobacco industry documents, containing more than 14 million documents, which was developed by a variety of tobacco companies and which relates to their advertising, marketing, manufacturing, sales, and research activities. For more information, see http://legacy.library.ucsf.edu (accessed March 2, 2015).

data from various sources, such as large-area administrative information and small-area detailed surveys. However, using such combinations would require considerable care. These challenges are compounded by the fact that, in general, "the data most useful for modeling is often among the most jealously guarded resources in many organizations" (North and Macal, 2007, p. 240).

A longer-term approach is to try to anticipate critical data needs and either fund or otherwise encourage the collection of data that best suit ABMs or other modeling approaches. Similarly, encouraging the standardization of data collection items and methods might improve model quality. Even for administrative data that are "routinely" collected, such as tobacco marketing and sales information or population smoking prevalence estimates, it could be possible to evaluate those data periodically for validity and consistency. It may also be possible to substitute existing or newly developed biomarkers of certain smoking behaviors for other forms of data collection, and in selected instances, information from other countries with similar populations may be of value.

Network data, which are thought to require the elucidation of an entire social network, are particularly difficult to collect. For example, many network measures, particularly centrality measures, are prone to biases (Costenbader and Valente, 2003; Kossinets, 2006; Smith and Moody, 2013). However, there may be some modeling efforts for which whole-network (sociometric) data are not required. As noted by the Statnet[7] Development Team, egocentric data was "long regarded as the poor country cousin in the network data family," yet such data "contain a remarkable amount of information" (Butts et al., 2014). Adding in egocentric (sampled) network questions would add information that is relevant to ABMs. In collecting network data of this type, one employs traditional survey methods to assemble representative samples of the population. Respondents could be asked questions about important contacts. For example: How many of your five best friends smoke? What are their relevant attributes? Do they know one another? How many of your family members smoke? Such questions could be added, for example, to the Behavioral Risk Factor Surveillance System or the TUS–CPS. Novel developments in sampled networks permit the simulation of disease outbreaks, which could be applied to behavioral "epidemics" as well as to infectious ones.[8]

The use of network-based ABMs in epidemiology has increased over the past decade. There are two issues in such modeling. One, ground truth

[7] Statnet is a statistical modeling package for the R platform. See http://statnet.org for more information (accessed March 2, 2015).

[8] For example, see details on the EpiModel at http://cran.r-project.org/web/packages/EpiModel/index.html for more information (accessed March 2, 2015).

(i.e., any data that capture the empirical process under investigation) is often limited to stylized facts and theoretical models with little empirical data. Two, confirmation rarely moves beyond internal validation and calibration. Networks provide structure—who interacts with whom—and are incorporated in a number of ways. The two most common approaches are to generate a stylized network or to input an actual network. In the case of generating a stylized network, the choice of which stylized network has implications for diffusion processes, including the time course and peak of an epidemic (Rahmandad and Sterman, 2008).

One of the key data needs for ABM is data that inform agent interactions, either with other agents or with the agent's environment. Such interactions are difficult or impossible to capture empirically. Traditional data from survey methods may not always provide the detailed data required for reproducing the relevant interactions, motives, sequence of events, or decision processes associated with the behavior of an agent. Alternative data collection methodologies could include qualitative methods (such as ethnography) that tap into the experience of social interactions (Falkin et al., 2007; Rothwell and Lamarque, 2011), experiential or situational sampling (e.g., ecological momentary assessment [EMA]; see Shiffman et al., 2008), and time-use data (e.g., the American Time Use Survey, or ATUS) that capture "with whom," "where," and "when" types of questions. The use of EMA has been particularly enlightening for understanding the context of tobacco cravings (Chandra et al., 2011). Time-use data capture a representative slice of daily activities; the ATUS sample is drawn from the Current Population Survey (CPS), which means it can be linked to the TUS–CPS (NCI, 2014). Such linkage provides a rich source of daily activities within a geographic context. Experimental or quasi-experimental data may also be relevant, such as those from random roommate assignments (Eisenberg et al., 2014), and data from randomized controlled trials of smoking cessation programs (Bullen et al., 2013; Strecher et al., 2008). Such studies, especially individual-level trials, would be useful in parameterizing empirically based rules for agent behavior.

Online platforms may offer yet another way to collect data. While tobacco companies are making extensive use of online social media to market their products, the tobacco control community is using online platforms to counter the marketing of tobacco products (Legacy, 2012), provide cessation support services and forums (Gutierrez and Newcombe, 2012), and mobilize advocates to strengthen tobacco-control efforts (Hefler et al., 2013). Because various stakeholders of the tobacco environment use online social media (see Box 6-1 for other online platforms and related trackers), enormous amounts of data have been generated, including the social connections and interactions among individuals online. Such data may be mined to better understand the diffusion dynamics of and the role of social network structure in tobacco use (Centola, 2013; Cobb et al., 2010). Content analysis is now possible on a massive scale, a development that could help enrich the understanding of the mechanisms that drive tobacco addiction (Myslín et al., 2013). It is important to keep in mind, however, that online and face-to-face networks are distinct and potentially interact with one another (Huang et al., 2014), so it will continue to be necessary to use a range of research methods.

There are also data needs at the aggregate (state/national/local) level. It will be necessary to remain vigilant in collecting both qualitative and quantitative observations of American tobacco use habits over time. Changes will likely occur in the types of tobacco user groups and their general characteristics, such as age, gender, race/ethnicity, socioeconomic status, cultural beliefs, health characteristics, the types of tobacco delivery devices used, and the use of other relevant substances. Of course, the policy questions themselves may change over time, which will also affect the nature of data collection. Not all of these changes can be easily predicted, making ongoing population tobacco surveillance necessary, if only for basic data needs and to identify more targeted surveys for policy promulgation. Such data would be useful as ground truth against which simulation results could be compared.

Other Types of Agents for Application in Agent-Based Models

Another area for future data collection is capturing information on the many agents that could be modeled in ABMs developed for tobacco control policy. The agents in SnapDragon (the central ABM evaluated in this report, see Chapter 5) are people and media, but other types of agents are possible. These agents could, for example, include state and local legislators, policy makers, and health departments if one wished to better understand how they approach tobacco control and regulation at the local level. Social networks, particularly the ways in which information and resources are shared among stakeholder groups in the tobacco control regulatory landscape, have been described by Luke and colleagues (Harris et al., 2008; Luke et al., 2010). Organizational collaborations among public health agencies, advocacy groups, and funders, among others, are critical in the dissemination of tobacco control research and evidence-based best practices (Luke et al., 2013). ABMs could also be used to consider the tobacco industry's behavior, with tobacco companies being the agents in the model. For example, an ABM could examine the role of current cigarette manufacturers in the alternative nicotine delivery market. ABMs that aim to capture industry behavior could complement other models and research in illuminating the implications of tobacco product use and could provide guidance on the type of industry data that is needed for policy evaluation. Agents may also include state excise tax collectors, private area-wide commerce organizations, and other organizations and government agencies that do not have tobacco regulation as their fundamental mission but whose policies impact tobacco use. For example, agents might include housing and environmental agencies, chambers of commerce, commercial trade organizations, or police organizations that may become involved in contraband tobacco products. Health systems and health professionals might also be considered as agents in some policy models.

Agents may also be "below the skin," as components of complex biological systems—for example, neurons, nicotine, nicotinic acetylcholine receptors, and cytochrome P450 enzymes could all be considered types of agents. Neural pathways have been identified as key components in the addiction process, which entails the activation of reward-learning circuits (Hyman et al., 2006; Koob and Le Moal, 2001). According to Hyman and colleagues, "Humans and animals rapidly learn cues and contexts that predict the availability of these 'addictive drugs'; once learned, these cues motivate drug seeking in humans and animal models" (Hyman et al., 2006, p. 567). Because addiction plays such a central role in tobacco use, modeling the process of addiction and the resulting difficult-to-change behavior could help strengthen ABMs.

The mathematical and computational modeling of biological systems has been helpful in understanding other disease processes, including hepatitis clearance and infectivity (Dahari et al., 2009), host–pathogen interactions (Stern et al., 2013), and inflammation and multiple organ failure (An, 2004, 2006). Relevant "below the skin" factors—that is, various elements that constitute and act on an organism's biological systems—have not been largely used in ABM of tobacco use. Biological factors are not necessarily required if they can be represented by simply using proxies or if the model is concerned primarily with above-the-skin factors. In other words, the level at which agents are specified will depend on the questions being asked of the model. Such low-level detail may be necessary if individual responses to nicotine (e.g., half-life) and the toxicity of a tobacco product are important, but they may be unnecessary or undesired if the model concerns diffusion of information or norms. Following the recommendations of a National Research Council report (2008), models should strive for parsimony and avoid "kitchen sink" approaches. The report noted that "models can become unwieldy when weighed down by a proliferation of features and variables" (p. 347). Nevertheless, given that tobacco use initiation and cessation are at least in part based on human physiology, the modeling of relevant biological mechanisms would need to at least be considered.

DATA COLLECTION AND MODEL DEVELOPMENT AT THE CENTER FOR TOBACCO PRODUCTS

In this report, the committee discussed the importance and challenges of incorporating data into ABMs that are intended to inform tobacco control policy. Models that use minimal amounts of data can be used to guide data collection and the development of future models. However, when the goal is to guide policies, data can help ensure that the agents capture realistic representations of actual entities to the extent possible, and data can also confirm the degree to which the model replicates or predicts real-world patterns, such as initiation and cessation processes among individuals or populations. In the tobacco control field, an assortment of data is available (as presented earlier in the chapter as well as in the tobacco use behavior section of Chapter 2), and much of these data could help strengthen ABMs developed to guide tobacco control policy. These data could be used creatively to inform models, and more data could be collected from efforts that go beyond traditional survey methods, such as gathering information from online social media platforms. Data collected with behavioral mechanisms in mind would allow agent-based modelers to capture more realistic agent characteristics as well as more realistic agent–agent and agent–environment interactions. It is important that these characteristics and interactions be captured meaningfully because they tend to be central elements of ABMs that aim to inform policy decisions, especially if the goal of the model is to understand how interdependent agent behavior will shape the outcomes experienced under a given policy (as discussed in Chapter 3). Because ABMs and other individual-level modeling techniques are promising tools to further our understanding of tobacco use behavior, it is worthwhile to collect such data. As a major funder and user of tobacco data (including for the modeling of tobacco use), CTP can help shape the tobacco data environment in the future.

Conclusion 6-1: The committee concludes that agent-based models designed to inform policy decisions require data on the underlying mechanisms governing behavior and on agent-to-agent and agent-to-environment interactions. Currently, these data are not commonly collected.

Recommendation 6-1: The Center for Tobacco Products should identify and help develop data sources relevant to the questions it is trying to address using agent-based models and other modeling approaches.

The use of data already being collected (either by CTP or other sources) could be incorporated into the modeling process. CTP could consider coordinating with other activities, such as the Tobacco Centers of Regulatory Science, to gather these data. As noted elsewhere in the report, models can help researchers identify data gaps and combine data from various sources (while recognizing the limitations of each), further guiding data collection and enhancing models used to inform policy.

To ensure that the processes of collecting the necessary data and of identifying agent attributes based on those data are done successfully, it is crucial to address implementation issues. Having the appropriate individuals overseeing these processes and ensuring that the models have broad input to inform them will both be important to the success of the models. Many different types of models have been developed by federal agencies; some of them developed within the agencies and others through contracts or grants. Regardless of where the models are developed, funders for policy-relevant models require access to expertise if they are to issue effective funding opportunity announcements or contracts; to make informed decisions about which modeling approaches are appropriate for the question at hand; to work effectively with the modeling team(s) throughout model development; to appropriately evaluate model inputs, processes, and outputs; and to interpret or translate model results appropriately to decision makers.

FDA is regularly confronted with uncertainty within the complex tobacco environment. Because of this, it will remain necessary to have models that represent potential tobacco policies to organize data, elucidate specific uncertainties, and forecast future scenarios. Because the use of models at CTP has the potential to affect regulatory decision making, it is essential that the development of these models be overseen by individuals who have the expertise and experience needed to maximize the benefit and reliability of the models. Subject-matter experts (that is, scientists and researchers who have a deep understanding of the tobacco literature and work in that field) could be essential partners in future CTP modeling endeavors (see Chapter 4 for more discussion on this topic).

> **Recommendation 6-2: The Center for Tobacco Products (CTP) should ensure that it has staff with, or access to, the necessary expertise to inform CTP's research, contracting, and evaluation efforts and to translate model results for various stakeholders.**

FDA could also consider obtaining input on the development of its models from tobacco stakeholders, including representatives from local, state, and federal public health agencies; scientists and other members of academia; other modelers; and end users, among others. CTP could acquire feedback in a number of ways, ranging from developing a standing expert panel to provide regular feedback on modeling initiatives, to using modeling networks or forums such as the Models of Infectious Disease Agent

Study (MIDAS),[9] the Cancer Intervention and Surveillance Modeling Network (CISNET),[10] the Drug Policy Modelling Program,[11] and the Energy Modeling Forum.[12,13]

Although individual models are a useful tool for informing policy decisions, having a range of modeling techniques will offer a fuller picture of the policy questions confronted by CTP—for example, by creating various models to approach the same question or process (e.g., multiple ABMs or ABMs and aggregate models), as is done by several modeling networks and forums.[14] The documentation of model inputs, activities, and outputs by the model developers (as discussed in Chapter 4) and a comparison of results with a rigorous discussion by the developers on why the results differ—or do not differ—will create a richer understanding of the models and the model results (Kuntz et al., 2013) and will help to address model uncertainty. Doing so will also help to increase policy makers' confidence in the model results, identify where assumptions need to be modified, and detect where further data are needed.[15]

Recommendation 6-3: The U.S. Food and Drug Administration should develop a range of models using various approaches. This would include agent-based models as well as other modeling approaches.

It is important to note that the range of models FDA could use includes not only those that FDA commissions or develops but also those that others have already developed or will develop to help guide tobacco control policy.

---

[9] For more on MIDAS, see http://www.nigms.nih.gov/Research/SpecificAreas/MIDAS/Pages/default.aspx (accessed March 2, 2015).

[10] For more on CISNET, see http://cisnet.cancer.gov (accessed March 2, 2015).

[11] For more on Drug Policy Modelling Program, see https://dpmp.unsw.edu.au (accessed March 2, 2015).

[12] For more on Energy Modeling Forum, see https://emf.stanford.edu (accessed March 2, 2015).

[13] Modeling networks and forums can take several forms but generally consist of a collaborative network of researchers who develop various types of models to understand the topic at hand. These models are often intended for policy makers, public health officials, and other researchers to help them make better-informed decisions on the topic of study (see Appendix A).

[14] For example, the modeling done by CISNET is collaborative, and members address a common question using a common dataset. For a description of a specific instance of this, see the July 2012 supplement of *Risk Analysis*, which was devoted to the CISNET modeling of smoking and lung cancer, and various CISNET collaborative articles on breast, colon and prostate cancer.

[15] See Appendix A for a discussion on the benefits of using a multiple model approach, using MIDAS as an example.

CONCLUSION

Although simulation modeling has been used for many years in tobacco control, CTP is still in the early stages of its efforts to use ABM to explore tobacco control policy and regulation. This report has illustrated many of the challenging and technical aspects surrounding ABMs. However, the committee believes that ABMs are a useful tool that could add to the understanding of tobacco use initiation, cessation, and relapse processes. The model developed for FDA (see Chapter 5) does not accurately represent many of the important characteristics of tobacco use, but there is much to be learned from its development that can be applied to future models of tobacco use, both agent-based and otherwise. There are some barriers to overcome, such as the collection of data to inform the development of ABMs and the elucidation of the empirical and theoretical challenges of specifying model inputs and appropriately interpreting model outputs (see Chapter 3). A strong evaluation framework (as described in Chapter 4) will be needed to track rigorous model development. As discussed in Chapters 3 and 4, it will be important to consult an interdisciplinary modeling team and subject-matter experts at the earliest stage of model conceptualization and then throughout the model development process in order to ensure that the model is grounded in the current state of tobacco science (that is, evidence-based research related to tobacco in the fields of epidemiology, social and behavioral sciences, biology, chemistry, and others), while carefully considering individual behavior. If the principles discussed in this report are followed, the value of ABMs for informing tobacco regulation will be greatly strengthened.

REFERENCES

Amerson, N. L., B. S. Arbise, N. K. Kelly, and E. Traore. 2014. Use of market research data by state chronic disease programs, Illinois, 2012–2014. *Preventing Chronic Disease* 11:E165.

An, G. 2004. In silico experiments of existing and hypothetical cytokine-directed clinical trials using agent-based modeling. *Critical Care Medicine* 32(10):2050–2060.

———. 2006. Concepts for developing a collaborative in silico model of the acute inflammatory response using agent-based modeling. *Journal of Critical Care* 21(1):105–110.

Apelberg, B. J., C. L. Backinger, and S. J. Curry. 2014. Enhancing youth tobacco surveillance to inform tobacco product regulation: Findings from the 2012 National Youth Tobacco Survey. *American Journal of Preventive Medicine* 47(2 Suppl 1):S1–S3.

Bero, L. 2003. Implications of the tobacco industry documents for public health and policy. *Annual Review of Public Health* 24(1):267–288.

Bullen, C., C. Howe, M. Laugesen, H. McRobbie, V. Parag, J. Williman, and N. Walker. 2013. Electronic cigarettes for smoking cessation: A randomised controlled trial. *Lancet* 382(9905):1629–1637.

Butts, C. T., M. Morris, P. N. Krivitsky, Z. Almquist, M. S. Handcock, D. R. Hunter, S. M. Goodreau, and S. B. de-Moll. 2014. Introduction to exponential-family random graph (ERG or p*) modeling with ergm. http://statnet.csde.washington.edu/workshops/UNBELT/current/ergm/ergm_tutorial.pdf (accessed November 7, 2014).

CDC (Centers for Disease Control and Prevention). 2014. *Surveillance and evaluation data resources for comprehensive tobacco control programs*. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

Centola, D. 2013. Social media and the science of health behavior. *Circulation* 127(21): 2135–2144.

Chandra, S., D. Scharf, and S. Shiffman. 2011. Within-day temporal patterns of smoking, withdrawal symptoms, and craving. *Drug and Alcohol Dependence* 117(2–3):118–125.

Cobb, N. K., A. L. Graham, and D. B. Abrams. 2010. Social network structure of a large online community for smoking cessation. *American Journal of Public Health* 100(7):1282–1289.

Costenbader, E., and T. W. Valente. 2003. The stability of centrality measures when networks are sampled. *Social Networks* 25(4):283–307.

CPC (Carolina Population Center, University of North Carolina at Chapel Hill). 1998. *National Longitudinal Study of Adolescent Health: Wave I*. http://www.cpc.unc.edu/projects/addhealth/codebooks/wave1 (accessed February 25, 2015).

Cruz, T. B. 2009. Monitoring the tobacco use epidemic IV. The vector: Tobacco industry data sources and recommendations for research and evaluation. *Preventive Medicine* 48(1 Suppl):S24–S34.

Cummings, K. M., C. P. Morley, J. K. Horan, C. Steger, and N. R. Leavell. 2002. Marketing to America's youth: Evidence from corporate documents. *Tobacco Control* 11(Suppl 1):I5–I17.

Dahari, H., E. Shudo, R. Ribeiro, and A. Perelson. 2009. Mathematical modeling of HCV infection and treatment. *Methods in Molecular Biology* 510:439–453.

Delnevo, C. D. 2014. Tobacco surveillance data for population modeling: Getting the inputs right. Paper presented at Tobacco Products Scientific Advisory Committee Meeting, Rockville, MD.

Eisenberg, D., E. Golberstein, and J. L. Whitlock. 2014. Peer effects on risky behaviors: New evidence from college roommate assignments. *Journal of Health Economics* 33:126–138.

Falkin, G. P., C. S. Fryer, and M. Mahadeo. 2007. Smoking cessation and stress among teenagers. *Qualitative Health Research* 17(6):812–823.

Farrelly, M. C. 2009. Monitoring the tobacco use epidemic V: The environment: Factors that influence tobacco use. *Preventive Medicine* 48(1 Suppl):S35–S43.

Gutierrez, K., and R. Newcombe. 2012. *Lessons learned globally: Tobacco control digital media campaigns*. Saint Paul, MN: Global Dialogue for Effective Stop Smoking Campaigns.

Haas, S., and D. Schaefer. 2014. With a little help from my friends? Asymmetrical social influence on adolescent smoking initiation and cessation. *Journal of Health and Social Behavior* 55(2):126–143.

Harris, J. K., D. A. Luke, R. C. Burke, and N. B. Mueller. 2008. Seeing the forest and the trees: Using network analysis to develop an organizational blueprint of state tobacco control systems. *Social Science and Medicine* 67(11):1669–1678.

Harris, K. M. 2013. *The Add Health study: Design and accomplishments*. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill.

Harris, K. M., C. T. Halpern, E. Whitsel, J. Hussey, J. Tabor, P. Entzel, and J. R. Udry. 2009. The National Logitudinal Study of Adolescent to Adult Health: Research design http://www.cpc.unc.edu/projects/addhealth/design (accessed March 2, 2015).

Hefler, M., B. Freeman, and S. Chapman. 2013. Tobacco control advocacy in the age of social media: Using Facebook, Twitter and change. *Tobacco Control* 22(3):210–214.

Huang, G. C., D. Soto, K. Fujimoto, and T. W. Valente. 2014. The interplay of friendship networks and social networking sites: Longitudinal analysis of selection and influence effects on adolescent smoking and alcohol use. *American Journal of Public Health* 104(8):e51–e59.

Hurt, R. D., S. A. Weston, J. O. Ebbert, S. M. McNallan, I. T. Croghan, D. R. Schroeder, and V. L. Roger. 2012. Myocardial infarction and sudden cardiac death in Olmsted County, Minnesota, before and after smoke-free workplace laws. *Archives of Internal Medicine* 172(21):1635–1641.

Hyman, S. E., R. C. Malenka, and E. J. Nestler. 2006. Neural mechanisms of addiction: The role of reward-related learning and memory. *Annual Review of Neuroscience* 29:565–598.

Koob, G. F., and M. Le Moal. 2001. Drug addiction, dysregulation of reward, and allostasis. *Neuropsychopharmacology* 24(2):97–129.

Kossinets, G. 2006. Effects of missing data in social networks. *Social Networks* 28(3):247–268.

Kuntz, K., F. Sainfort, M. Butler, B. Taylor, S. Kulasingam, S. Gregory, E. Mann, J. M. Anderson, and R. L. Kane. 2013. *Decision and simulation modeling in systematic reviews*. Rockville, MD: Agency for Healthcare Research and Quality.

Lakon, C. M., C. Wang, C. T. Butts, R. Jose, D. S. Timberlake, and J. R. Hipp. 2014. A dynamic model of adolescent friendship networks, parental influences, and smoking. *Journal of Youth and Adolescents*.

Legacy. 2012. *Truth health fact sheet*. http://www.legacyforhealth.org/content/download/621/7337/file/truth_fact_sheet_January_2012.pdf (accessed February 18, 2015).

Luke, D. A., J. K. Harris, S. Shelton, P. Allen, B. J. Carothers, and N. B. Mueller. 2010. Systems analysis of collaboration in 5 national tobacco control networks. *American Journal of Public Health* 100(7):1290–1297.

Luke, D. A., L. M. Wald, B. J. Carothers, L. E. Bach, and J. K. Harris. 2013. Network influences on dissemination of evidence-based guidelines in state tobacco control programs. *Health Education and Behavior* 40(1 Suppl):33S–42S.

Mercken, L., T. A. Snijders, C. Steglich, and H. de Vries. 2009. Dynamics of adolescent friendship networks and smoking behavior: Social network analyses in six European countries. *Social Science and Medicine* 69(10):1506–1514.

Mermelstein, R. J. 2014. Adapting to a changing tobacco landscape: Research implications for understanding and reducing youth tobacco use. *American Journal of Preventive Medicine* 47(2 Suppl 1):S87–S89.

Moore, T. W., P. D. Finley, N. S. Brodsky, T. J. Brown, B. Apelberg, B. Ambrose, R. J. Glass. In press a. Modeling education and advertising with opinion dynamics. *The Journal of Artificial Societies and Social Simulation*.

Moore, T. W., P. D. Finley, B. J. Apelberg, B. Ambrose, N. S. Brodsky, T. J. Brown, C. Husten, R. J. Glass. In press b. An opinion-driven behavioral dynamics model for addictive behaviors. *European Physical Journal B*.

Myslín, M., S. H. Zhu, W. Chapman, and M. Conway. 2013. Using Twitter to examine smoking behavior and perceptions of emerging tobacco products. *Journal of Medical Internet Research* 15(8):e174.

NCI (National Cancer Institute). 2014. *Tobacco use supplement—current population survey fact sheet*. http://appliedresearch.cancer.gov/tus-cps/TUS-CPS_fact_sheet.pdf (accessed November 8, 2014).

The Neilsen Company. 2014. *About us*. http://www.nielsen.com/us/en/about-us.html (accessed November 10, 2014).

North, M. J., and C. M. Macal. 2007. *Managing business complexity: Discovering strategic solutions with agent-based modeling and simulation*. Oxford, UK: Oxford University Press.

NRC (National Research Council). 2008. *Behavioral modeling and simulation: From individuals to societies*. Washington, DC: The National Academies Press.

NYSDOH (New York State Department of Health). 2011. *Key tobacco control outcome indicators*. New York State Department of Health. https://www.health.ny.gov/prevention/tobacco_control/docs/2011-09_key_tobacco_outcome_indicators.pdf (accessed October 28, 2014).

PATH (Population Assessment of Tobacco and Health). 2015a. *Study overview*. http://www.pathstudyinfo.nih.gov/UI/StudyOverviewMobile.aspx (accessed February 18, 2015).

———. 2015b. *FAQS for researchers*. http://www.pathstudyinfo.nih.gov/UI/FAQsResMobile.aspx (accessed November 21, 2014).

Pollard, M. S., J. S. Tucker, H. D. Green, D. Kennedy, and M. H. Go. 2010. Friendship networks and trajectories of adolescent tobacco use. *Addictive Behaviors* 35(7):678–685.

Rahmandad, H., and J. Sterman. 2008. Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models. *Management Science* 54(5):998–1014.

Rothwell, E., and J. Lamarque. 2011. The use of focus groups to compare tobacco attitudes and behaviors between youth in urban and rural settings. *Health Promotion Practice* 12(4):551–560.

Schaefer, D. R., S. A. Haas, and N. J. Bishop. 2012. A dynamic model of U.S. adolescents' smoking and friendship networks. *American Journal of Public Health* 102(6):e12–e18.

Schaefer, D. R., J. Adams, and S. A. Haas. 2013. Social networks and smoking: Exploring the effects of peer influence and smoker popularity through simulations. *Health Education and Behavior* 40(1 Suppl):24S–32S.

Shiffman, S., A. A. Stone, and M. R. Hufford. 2008. Ecological momentary assessment. *Annual Review of Clinical Psychology* 4:1–32.

Smith, J. A., and J. Moody. 2013. Structural effects of network sampling coverage I: Nodes missing at random. *Social Networks* 35(4):652–668.

Snijders, T. A., G. G. Van de Bunt, and C. E. Steglich. 2010. Introduction to stochastic actor-based models for network dynamics. *Social Networks* 32(1):44–60.

Stern, J. R., A. D. Olivas, V. Valuckaite, O. Zaborina, J. C. Alverdy, and G. An. 2013. Agent-based model of epithelial host-pathogen interactions in anastomotic leak. *Journal of Surgical Research* 184(2):730–738.

Strecher, V. J., J. B. McClure, G. L. Alexander, B. Chakraborty, V. N. Nair, J. M. Konkel, S. M. Greene, L. M. Collins, C. C. Carlier, C. J. Wiese, R. J. Little, C. S. Pomerleau, and O. F. Pomerleau. 2008. Web-based smoking-cessation programs: Results of a randomized trial. *American Journal of Preventive Medicine* 34(5):373–381.

Terry-McElrath, Y. M., S. Emery, M. A. Wakefield, P. M. O'Malley, G. Szczypka, and L. D. Johnston. 2011. Effects of tobacco-related media campaigns on smoking among 20–30-year-old adults: Longitudinal data from the USA. *Tobacco Control* 22(1):38–45.

Valente, T. W., J. B. Unger, A. Ritt-Olson, S. Y. Cen, and C. Anderson Johnson. 2006. The interaction of curriculum type and implementation method on 1-year smoking outcomes in a school-based prevention program. *Health Education Research* 21(3):315–324.

Valente, T. W., K. Fujimoto, D. Soto, A. Ritt-Olson, and J. B. Unger. 2013. A comparison of peer influence measures as predictors of smoking among predominately Hispanic/Latino high school adolescents. *Journal of Adolescent Health* 52(3):358–364.

Appendix A

Considerations and Best Practices in Agent-Based Modeling to Inform Policy

Ross A. Hammond

1. INTRODUCTION

Agent-based modeling (ABM) is a powerful tool that is being used to inform policy or decisions in many fields of practical importance. Recent examples include land-use and agricultural policy (Berger et al., 2007; Berger and Troost, 2014; Brady et al., 2012; Guzy et al., 2008; Happe et al., 2008; Happe et al., 2006; Heckbert, 2011), ecosystem and natural-resource management (Heckbert et al., 2010; Schlüter and Pahl-Wostl, 2007), control of communicable disease outbreaks (Burke et al., 2006; Epstein, 2004; Epstein, 2009; Eubank et al., 2004; Ferguson et al., 2006; Germann et al., 2006; Lee et al., 2010; Longini et al., 2005; Longini Jr et al., 2007; Yang et al., 2009), marketing (Garcia and Jager, 2011; Rand and Rust, 2011) and private-sector logistics and strategy (Frederick, 2013; North et al., 2010; Rand and Rust, 2011), economic policy (Dawid and Fagiolo, 2008; Frederick, 2013; LeBaron and Winkler, 2008), electoral design (Bendor et al., 2003; Laver, 2005), and education (Maroulis, 2014).

In this paper, I discuss some of the features of ABM that make it compelling for such purposes (especially in the context of public health), lay out the process and challenges involved in using ABM, and offer some important best practices for rigorous and effective use. This is not a textbook or a how-to manual; it is intended as an overview of the major topics and considerations involved in the use of ABM for policy. It just scratches the surface in most cases but provides some references for further reading. I will argue that the use of ABM to inform policy making or decision making can be both promising and practical but is often challenging and requires great care in practice.

1.1 What Is Agent-Based Modeling?

Agent-based computational modeling (ABM) is an approach to modeling complex social dynamics that has developed in recent decades, facilitated by increased computational power. In an ABM, actors in a system are represented as autonomous individuals in a computer program. They are given rules that govern their behavior, including adaptation and interaction with each other and with their environment through time, and a starting configuration. The ABM then simulates[1] both individual trajectories and population-level patterns or outcomes, which are generated from the bottom up by the decentralized actions and interactions of the agents. Such a model provides mechanistic mapping from individual-level assumptions to coevolving population-level dynamics. Assumptions can be informed by data or theory, and outcomes at both the individual and population levels can be compared with data statistically. ABM allows enormous flexibility in assumptions, and agents can be modeled at any level (or multiple levels) of scale.

1.2 Why Agent-Based Modeling?

Like other modeling methods, this technique has both advantages and important limitations. The particular advantages of ABM come from its flexibility, which can help model designers and users to manage three particular challenges that complexity poses for researchers and policy makers alike: heterogeneity, spatial structure, and adaptation.

Heterogeneity

Real-world complex systems are often characterized by substantial heterogeneity among individuals. Among individuals of a particular type, this might include biological diversity (e.g., in genes, microbiome, sensitivity to reward), behavioral diversity (e.g., in decision-making, psychology, personality), demographic diversity (e.g., in socioeconomic status, race, sex, and age) or diversity in context or prior experiences (see "Spatial Structure" on next page). There may also be substantial heterogeneity in *types of actors* that are important in a system's behavior; for example, the outcome of childhood obesity is driven partly by such diverse actors as parents, community stakeholders, school employees, health professionals, food companies, and the children themselves. Types of actors may differ substantially in information sets, goals, incentive structure, constraints, and so on.

[1] ABM simulations typically involve stochastic elements to represent phenomena that may be either inherently unpredictable (for example, everyone with whom an agent will come into contact on a given day) or about which the researcher cannot make precise assumptions (such as the explicit structure of a complex network).

By explicitly modeling every individual actor (within the model boundary), ABM allows rich representation of heterogeneity. No aggregation (such as "representative agents," compartments, or mean-field approximations) is required in an ABM, although aggregation can be accommodated if useful. Taking heterogeneity into account can be critical in the design of successful interventions into complex systems (IOM, 2012; Mabry et al., 2010; Sterman, 2006).

Spatial Structure

An important advantage of ABM is the ability to include structurally rich, dynamic, and heterogeneous representations of social or environmental exposures and influences. For example, ABM can incorporate explicit representations of geography from GIS data (Axtell et al., 2002; Brown et al., 2005a; Brown et al., 2005b; Magliocca et al., 2014; Page, 1999; Sun et al., 2014) or detailed social network structures (Hammond and Ornstein, 2014; Zhang et al., 2014)—representations that may be difficult in standard analytical approaches (Axelrod et al., 2004; Eubank et al., 2004; Page, 1999), which tend to rely on mean-field or other approximations. By directly incorporating sophisticated spatial elements, ABM can effectively model dynamics that result from exposures across space and time (such as advertising or air pollution exposure), patterns of contact between individuals (central to epidemic spread or social influence through networks), the impact of context on decision making, and geographic constraints on choice set (such as the distribution of retailers with heterogeneous characteristics).

ABM not only allows incorporation of spatial elements that affect agents and their interaction with one another, but it also allows modeling of the *coevolution* of environment and individual behavior, on potentially divergent time scales: for example, the coevolution of retail geography and consumer purchases or of individual choices and social norms (see below).

Adaptation and Coevolution (Potentially Across Scales)

The ABM technique is particularly adept at modeling interaction and adaptation. By modeling at the individual level, ABM allows consideration of multiple interdependent factors that influence an outcome (such as health status). Because ABMs are dynamic, individual-level *adaptation* can also be represented, whether it takes the form of biological adaptation (as in an addiction process or physiological changes due to weight gain) or of behavioral adaptation (as in learning). A dynamic, individual-level focus also allows ABM to consider such phenomena as path-dependence (Page, 2006), which is important for life-course models that focus on key development windows or accumulation of exposures.

By modeling *populations* of individuals, ABM can also capture the *interaction* of actors with each other and with their coevolving environments. This type of interaction and feedback between individual and social levels of scale is important for the study of such phenomena as interacting social influence and social selection processes in adolescents, strategic coevolution of pro-tobacco and anti-tobacco marketing, and the bidirectional influence of social norms and individual behavior.

ABM is also well positioned to study mechanisms or pathways that cross multiple levels of scale. Agents themselves may be modeled on different levels of scale, for example, "employee" agents and the "corporation" agents for whom they work. In addition, ABM offers the opportunity to embed rich depictions of mechanisms within an agent (e.g., physiology or neurobiology) that take as inputs factors outside the agent (e.g., environmental exposure to food or marketing) and interact with between-agent dynamics (e.g., social norms). This enables ABM to consider topics in public health that cross the "skin barrier," for example (Glass and McAtee, 2006; Hall et al., 2014; Hammond, 2009; Hammond and Ornstein, 2014; Hammond et al., 2012; Mabry et al., 2010).

Policy Resistance

Heterogeneity, spatial structure, and adaptation all complicate analysis, and many analytical approaches struggle to address one or more of these features. The presence of these characteristics in a system may also contribute to *policy resistance* (Sterman, 2006). Anticipating adaptive (and potentially diverse) responses of a system to an intervention can be critical in designing effective policies. Behavioral and biological adaptation by individual actors can change the impact of an intervention for better or for worse. Interventions that appear promising on a small scale (or in one part of a system) may also run into "equilibrium dilution" or even produce net negative effects as adaptive individual or organizational responses on larger scales (or elsewhere in the system) come into play. The flexibility of ABM in capturing adaptation and heterogeneity thus makes it a potentially useful tool to inform decision-making in complex systems.

1.3 A Brief Overview of Agent-Based Modeling in Various Fields

Application of ABM first proliferated in biology and in social science and initially focused on important uses of the technique for theory and hypothesis development. Early examples in social science included work on cooperation (Axelrod, 1997b), electoral and bureaucratic dynamics (Bendor et al., 2003; Bendor and Moe, 1985; Kollman et al., 1992, 1997; Laver, 2005), conflict (Bhavnani and Miodownik, 2009; Epstein, 2002), and segregation (Bruch and Mare, 2006; Schelling, 1971; Xie and Zhou, 2012). Early work using ABM in evolutionary biology (Axelrod et al., 2004; Hammond and Axelrod, 2006a; Holland, 1992; Nowak, 2006; Ohtsuki et al., 2006) and ecology (DeAngelis and Mooij, 2005; Heckbert et al., 2010) also contributed in important ways to theory development. Many of these efforts leveraged the ability of ABM to capture heterogeneity, spatial structure, and adaptation to generate important new insights. With growing computing power, application of ABM expanded to such fields as education (Maroulis et al., 2014), anthropology (Axtell et al., 2002), economics and finance (Dawid and Fagiolo, 2008; Dawid and Neugart, 2011; Farmer, 2000; Farmer and Foley, 2009; LeBaron and Winkler, 2008; Tesfatsion and Judd, 2006), marketing (North et al., 2010; Rand and Rust, 2011), and land use (Berger et al., 2007; Berger and Troost, 2014; Brady et al., 2012; Brown et al., 2005a; Brown et al., 2005b; Guzy et al., 2008; Happe et al., 2008; Happe et al., 2006; Heckbert, 2011; Magliocca et al., 2014; Sun et al., 2014). ABM also began to be applied in a broader set of ways, including models engaged with large data sets (Axtell et al., 2002; Bruch and Mare, 2006; Farmer and Foley, 2009) and models designed to engage with or inform policy and to address policy resistance (Berger et al., 2007; Brown et al., 2005a; Brown et al., 2005b; Dawid and Fagiolo, 2008; Farmer, 2000; Guzy et al., 2008; Happe et al., 2008; Happe et al., 2006; Heckbert, 2011; LeBaron and Winkler, 2008; Magliocca et al., 2014; Schlüter and Pahl-Wostl, 2007; Sun et al., 2014).

A very recent, but rapidly growing, application area for ABM is in public health. Initial applications of ABM to public health focused on the epidemiology and control of communicable disease (Burke et al., 2006; Epstein, 2004; Epstein, 2009; Eubank et al., 2004; Ferguson et al., 2006; Germann et al., 2006; Lee et al., 2010; Longini et al., 2005; Longini Jr et al., 2007; Yang et al., 2009). A large network of modelers (MIDAS[2]) funded by the National Institutes of Health (NIH) has had substantial scientific and policy impact using ABM among other modeling approaches (see section 3.1 below). The last 5 years have seen growing recognition of the potential for ABM to yield new insights on a wide array of topics in public health, particularly in light of the importance of heterogeneity, spatial structure, and adaptation that have been informed by other fields (Brown et al., 2005a; Brown et al., 2005b; Magliocca et al., 2014; Sun et al., 2014), and the approach was highlighted in three recent Institute of Medicine reports (IOM, 2010, 2012, 2013). This broader awareness has led to the recent proliferation of work, including the creation of two additional NIH-funded model-

[2] For more information, see http://www.nigms.nih.gov/Research/SpecificAreas/MIDAS/Pages/default.aspx.

ing networks that use ABM: one focused on obesity (NCCOR Envision[3]) and one on health disparities (NICH[4]). Initial ABM studies in these areas include (COSSA, 2014; Hall et al., 2014; Hammond, 2009; Hammond and Ornstein, 2014; Hammond et al., 2012; Zhang et al., 2014).

2. MANY (DISTINCT) USES FOR AGENT-BASED MODELING

The overview above highlighted the growing array of topics to which ABM is being applied, but also began to draw out several distinct *ways* in which the technique can be used. Models in general (and ABM in particular) can be used for a variety of specific purposes as part of a research, education, or decision-support agenda (Epstein, 2008). Four especially common uses of ABM are (1) formulating or testing explanatory hypotheses about (potentially unobservable) mechanisms driving observed patterns in the real world, (2) bridging individual-level assumptions and population-level dynamics, (3) guiding data collection or empirical analysis by pinpointing especially important gaps or by discovering new questions, and (4) informing the design or evaluation of interventions (including policy choices).

In each of these uses, ABM can yield compelling insights to complement existing approaches—although the particular perspective that it provides is not always well suited for every topic or question (see Heckbert et al., 2010, and others for "litmus tests" of suitability). In the rest of this paper, I will focus on the specific use of ABM as a decision-support tool to inform policy or intervention design and evaluation.

2.1 Policy as a Specific Use for Agent-Based Modeling

Computational or mathematical models (including ABM) offer a number of potential advantages to a decision maker. By making explicit the assumptions, key pathways, and uncertainties involved (along with the mapping of all three of these onto potential outcomes), models can help decision-makers to revisit and discuss implicit mental models that may be driving the decision process. Explicit models are more easily tested, both for internal consistency and for external fidelity. Models can also be especially useful tools when fielding real-world experiments to inform policy choice is difficult, overly expensive, time-consuming, unethical, or impractical. An additional advantage offered by models such as ABM lies in their ability to uncover potentially unanticipated adaptive system responses that a policy or intervention might trigger (see section 1.2). ABM can also help a decision maker understand the implications heterogeneity (across individuals, con-

[3] For more information, see http://www.nccor.org/envision/index.
[4] For more information, see http://sitemaker.umich.edu/nich/about_nich.

texts, or time) may have for the impact of a policy in the longer term or in contexts differing from those for which empirical evidence is available. Finally, models can sometimes be of particular use when processes of policy implementation or even policy making itself are the focus.

2.2 Three Specific "Modalities" for Informing Policy with Agent-Based Modeling

The use of ABM to inform policy or decision making comes with its own particular set of considerations. ABMs that inform policy fall into three distinct categories: prospective policy models, retrospective policy models, and indirect policy models.

Prospective policy models (also sometimes called *ex ante* models) help to inform the design of policies or interventions by elucidating their potential effects. Such models contain representations of key dynamic mechanisms in a system, along with explicit representations of one or more policy choices, and they allow comparison of policy options within the simulated system. This process can aid in the design of policies or interventions by:

- Identifying leverage points where small shifts induced by targeted policies can generate large shifts in systemic outcomes or dynamics (such as "tipping points"). This may help to identify previously unnoticed opportunities or strategies for intervention.
- Elucidating potential linkages (trade-offs or synergy) between multiple policies or intervention elements in a complex system. This may help to facilitate coordination across "silos" in government or society, as needed for "systems" interventions (Huang et al., 2011; Nader et al., 2012).
- Allowing experimentation "in silico" to understand full potential consequences (intended or unintended) of interventions, which may include counterintuitive or unexpected impacts. This is of particular use when "in vivo" or "in vitro" experimentation (for example, through a randomized clinical trial) is not practical.
- Anticipating a variety of possible future scenarios that may unfold, incorporating both uncertainty and policy choices, or helping to elucidate how an intervention design might "scale," translate to a novel context, or play out in the long term.

Models (including ABM) are most effective as one input into a multi-faceted decision-making process; they generally cannot eliminate uncertainty or the need for judgment in weighing difficult trade-offs. They can, however, be of substantial help to decision makers in managing both complexity and uncertainty.

Retrospective policy models help to understand the underlying reasons for (retrospectively observed) success or failure of a policy or intervention that is already in place. To do this, they leverage the ability of ABM to provide insight into complex and dynamic mechanisms that are at work in a system which may not be directly observable (see section 1.2 above). In some settings, data may not exist (or might even be impossible to collect) to disentangle multiple simultaneously occurring mechanisms. ABM can help with causal inference in such circumstances. In the context of an intervention evaluation, this type of model can help evaluators to understand *why* and *how* elements of the intervention may have succeeded or failed. By facilitating consideration of heterogeneity (see section 1.2), ABM can also help to understand differential success of a policy or intervention across subpopulations or contexts. This may be critical for consideration of scaling and translation of successful interventions. In practice, retrospective modeling may often be combined with subsequent prospective modeling that leverages lessons learned from existing data to design improved interventions.

Indirect influence on policy or decision making may also come from models that are not explicitly aimed at consideration of policy choices. ABM offers extensive capabilities for understanding etiology, bidirectional relationships between system structure and individual behavior over time, and the operation of pathways that cross levels of scale. This type of model generally does not contain any explicit representation of policies or interventions and thus does not directly simulate the potential impacts of policy choices. Nonetheless, discoveries derived from this type of model may have important implications for policy—including identification of key leverage points, mechanisms, or windows of opportunity for intervention. Application of such insights within a policy-making process must be done with care and may require further simulation modeling that explicitly contains representation of the policy choices under consideration.

3. ILLUSTRATIVE EXAMPLES OF POLICY-RELEVANT AGENT-BASED MODELING

In this section, I provide brief descriptions of models that illustrate how policy may be informed by ABM in each of the three ways described above. The examples chosen are focused on public health where possible but also include a sampling of work from the social sciences; the set of examples here is by no means comprehensive.

3.1 Examples of "Prospective" Agent-Based Modeling to Inform Policy Design

Models of Infectious Disease

One of the earliest applications of ABM in public health has been in the modeling of communicable disease, and much of this work has had an explicit prospective focus on policy or intervention design. In 2003, the NIH National Institute of General Medical Sciences formed a collaborative network of scientists who were using modeling to understand infectious disease dynamics (MIDAS). The network, which now includes almost 100 scientists, has helped to pioneer the use of computational models (including ABM) to inform policies aimed at preparation for or response to epidemics. MIDAS has generated numerous scientific advances (Burke et al., 2006; Epstein, 2004; Epstein, 2009; Eubank et al., 2004; Ferguson et al., 2006; Germann et al., 2006; Lee et al., 2010; Longini et al., 2005; Longini Jr et al., 2007; Yang et al., 2009) and received the Distinguished Service Award from the U.S. Department of Health and Human Services for contributions to policy. The use of ABM in MIDAS has included both small-scale or exploratory models (Epstein, 2004; Epstein et al., 2008) and large-scale ones (Epstein, 2009; Eubank et al., 2004), with complex models built up in layers through iteration with exploratory and empirical work over a number of years (see section 4.3, BP3 below). These models leverage ABM's ability (see section 1.2) to include more realistic mixing patterns (explicit geography and/or networks), extensive heterogeneity (demographic, immunological, or behavioral), and adaptive behavior change by individuals in response to epidemics or to intervention elements (for example, protective self-isolation or decisions about care-seeking or vaccine acceptance). Some models cross many levels of scale from biological (disease progression and host-response within an individual person or virus evolution) to global (air travel or vaccine production). The experience of MIDAS has also helped to elucidate best practices for communicating models to policy makers (see section 4.3) and has underlined the value of multiple methods and multiple models in increasing confidence in policy-oriented findings (also see section 4.3).

One early MIDAS model that provides clear illustration of the prospective use of ABM to inform policy design can be found in work on smallpox preparedness (Burke et al., 2006; Epstein, 2004; Longini Jr et al., 2007). This model began with a stylized representation of individual movement across key social contexts identified in previous epidemiological work—agents in the model move between and spend time in households, workplaces or schools, and hospitals—and drew on appropriate demographic data. An "index case" of smallpox was introduced into this artificial population, and the spread of the pathogen (with characteristics drawn from empirical evidence on natural history) was simulated. The model was then used as a "virtual laboratory" to allow experimentation with varying policies for containment of the epidemic through vaccination. The model allowed comparison of potential impacts of policies already under consideration, but it also made use of the individual-level dynamic data created by the simulations (which provided a detailed account of how smallpox spread through a community) to identify *novel* policy options that focus on particularly high-leverage intervention targets to allow maximum effectiveness with minimal vaccine use.

The use of this type of model is of particular importance for policy discussions surrounding potential responses to bioterror—a circumstance that does not lend itself to real-world experimentation but would demand well-articulated and rapid policy response. By allowing prospective consideration of options *in silico*, making use of the best available data and capturing the inherent uncertainties (e.g., inexact pathogen parameters, timing and location of early cases) the models can be a key input into planning and decision-making. ABM on a much larger scale also proved useful in assisting policy response to the emerging H1N1 influenza epidemic of 2009–2010 (Epstein, 2009; Eubank et al., 2004; Ferguson et al., 2006; Germann et al., 2006; Lee et al., 2010; Longini et al., 2005; Yang et al., 2009). In that case, models had to address more extensive variation both in geographic context (from emergence of the virus in Asia to its spread around the globe to the United States) and in potential policy options (from antiviral prophylaxis to school closure to quarantine), but benefited from ongoing surveillance as the early epidemic unfolded. Earlier work in MIDAS helped to make possible the development and deployment of sophisticated models that were needed to inform policy response at both national and regional levels during the crisis. As in the case of smallpox, a primary use of ABM was for prospective consideration of varying mixtures of policy options across various contexts, with a clearly defined objective of effectively containing the epidemic.

Other Exploratory Work in Public Health

The use of ABM as a tool for prospective consideration of policy options in a public health context has begun to spread outside of infectious disease, including work in disaster preparedness (Epstein et al., 2011). In tobacco control, early development work for this type of ABM is under way. One example is the *Tobacco Town* project (Luke et al., 2014), which leverages the flexibility of ABM in representing detailed geography (see section 1.2) to consider tobacco control policies that are inherently spatial in nature (such as point-of-sale policies). This effort draws on demographic data, travel data from ecological momentary assessment, and retail exposure and purchase data to simulate representative communities in which point-of-sale policies might be deployed (FDA, 2013). The models include consideration of adaptive individual responses to environmental changes. The goal is to provide an *in silico* policy laboratory to understand the potential effects (intended or unintended) of retailer-based policy options such as zoning, licensing, and type-specific retailer density reduction across a variety of contexts and over both the short term and the longer term.

Outside of Public Health

ABM has been used extensively outside of public health as a tool for prospectively informing policy or interventions, including work on retirement policy (Axtell and Epstein, 1999), anticorruption interventions (Hammond, 2008), and agricultural production policies (Berger and Troost, 2014). This type of application for ABM has also appeared in the private sector (e.g., for consideration of changes to logistics, marketing, or strategy) (Frederick, 2013; North et al., 2010; Rand and Rust, 2011). Agents in these models represent (*inter alia*) current or potential retirees, bureaucrats, farmers, landowners, consumers, and employees.

3.2 Examples of "Retrospective" Agent-Based Modeling

Retrospective use of ABM to understand differential success of policies and interventions in public health has only recently begun to emerge,[5] but examples of this use are more widespread in social science. One illustrative example comes from political science consideration of real-world electoral systems and their implications for party competition and bureaucratic politics (Laver, 2005; Laver and Sergenti, 2011). In this work, agent-based models of multiparty competition (in which political party leaders and voters are types of agents) are applied to understand the historical trajectories of party policies and vote shares in 10 European countries. Another example is the recent use of ABM to study the economics of systemic risk in the housing market (Geanakoplos et al., 2012). This work looks retrospectively at policies that were in place during the housing boom and bust of 1997–2009 and develops an ABM of the underlying mechanisms (individual-level incentives and behavioral adaptations to the policies) that produced the observed outcome.

[5] For an example, see http://compactstudy.weebly.com.

3.3 Examples of "Indirect" Policy Implications from Agent-Based Modeling Focused on Mechanisms

Early applications of ABM in obesity research have focused on elucidating complex etiology. Obesity results from a multiscale system, with behaviors and outcomes driven by interacting mechanisms that sometimes cross levels of scale. ABM has the potential to offer new insights into mechanisms and to connect research focused on "below the skin" with that focused on "above the skin" (Hammond, 2009). Some of these models are beginning to offer insights that may have important *indirect* policy implications. For example, recent work focused on understanding preference formation builds on existing evidence in neuroscience on key brain systems that are involved in controlling eating behavior (Hall et al., 2014), making use of ABMs' ability to cross the "skin barrier" by embedding dynamic reward-learning processes within agents while exposing agents to different external sequences of environmental food exposures. The resulting model (Hammond et al., 2012) illustrates how early food exposures can strongly shape food preferences in ways that have substantial inertia in the face of subsequent changes in food opportunities; the model also shows how preference formation can be *path dependent* in the sequence of experience exposures (see section 1.2). Although this model does not directly consider any specific real-world policy, it offers potential implications for both targeting and timing of interventions to prevent obesity by encouraging the formation of preferences for healthy food. Another group of recent ABM papers focuses on the role of social networks and social influence in obesity, elucidating potential dynamic mechanisms through which social influence occurs (and which may potentially be harnessed for interventions) (Bahr et al., 2009; Hammond and Ornstein, 2014; Zhang et al., 2014). Some of these models include empirically grounded biological processes (Hall, 2010) that interact with the social level (Hammond and Ornstein, 2014); others begin to explore modalities for interventions to harness social forces while stopping short of prospectively modeling any specific real-world policy in detail (Bahr et al., 2009; Zhang et al., 2014).

In social science, examples of ABM with important indirect implications for policy include canonical work on the drivers of segregation (Bruch and Mare, 2006; Xie and Zhou, 2012) and work on the underlying mechanisms that may explain the ubiquity of ethnocentrism (Hammond and Axelrod, 2006b). In both cases, the models do not directly simulate specific policy or intervention choices—but they elucidate powerful pathways (and sometimes specific levers in the form of key variables) that could be harnessed for policy or intervention purposes (see, for example, Axelrod, 2004, which discusses potential application of insights from the ethnocentrism model to security issues of central Asia).

4. USING AGENT-BASED MODELING: "UNDER THE HOOD" AND BEST PRACTICES

As described in the previous sections of this paper, ABM can provide a powerful, flexible tool with high potential to offer meaningful insights for both scientific research and policy design. The very power and flexibility that make ABM appealing can also make it challenging to use appropriately, however. Section 1.2 (and 4.1 below) make clear that ABM involves many distinct choices about implementation, as well as many assumptions in translation from the real world into the computational world. Like any modeling technique, results from ABM flow directly from the inputs; thus the conclusions reached are only as strong as the inputs on which they are based. The specificity of an ABM (its ability to elucidate very specific operationalizations of mechanisms and actors) is part of its power; but it involves similar challenges, with great care needed in generalizing from the conclusions of a particular model. Because ABM is also a relatively new technique, opportunities for formal training and available reference materials, such as textbooks, are still limited. For all of these reasons, attention to emergent best practices in ABM is of particular importance.

This section of the paper begins by laying out the key elements and steps that go into constructing and using an ABM and then lays out a number of best practices for each of the steps in the process.

4.1 Key Building Blocks of an Agent-Based Model

Although ABMs are quite diverse as a group, reflecting diversity in topic (see section 1.3) and goal (see section 2), they share a set of fundamental building blocks. Clearly specifying and articulating these core pieces are critical for all stages of model construction, use, and communication (see section 4.2)—and understanding the number of design choices implicit in these building blocks helps to motivate some of the best practices discussed below (see section 4.3).

The elements of an ABM may be organized according to the "**PARTE**" framework: **P**roperties, **A**ctions, **R**ules, **T**ime, and **E**nvironment. The first three elements (Properties, Actions, Rules) define the agents, while the next two (Time and Environment) define the context (see Figure A-1).

Properties are characteristics of individual agents (such as sex, age, disease state, wealth, and body mass index). An agent property can be:

- *Mutable or immutable over time* (within the simulation). This is a design choice: a model focusing on a single school year might treat age as immutable, whereas allowing age to change within the

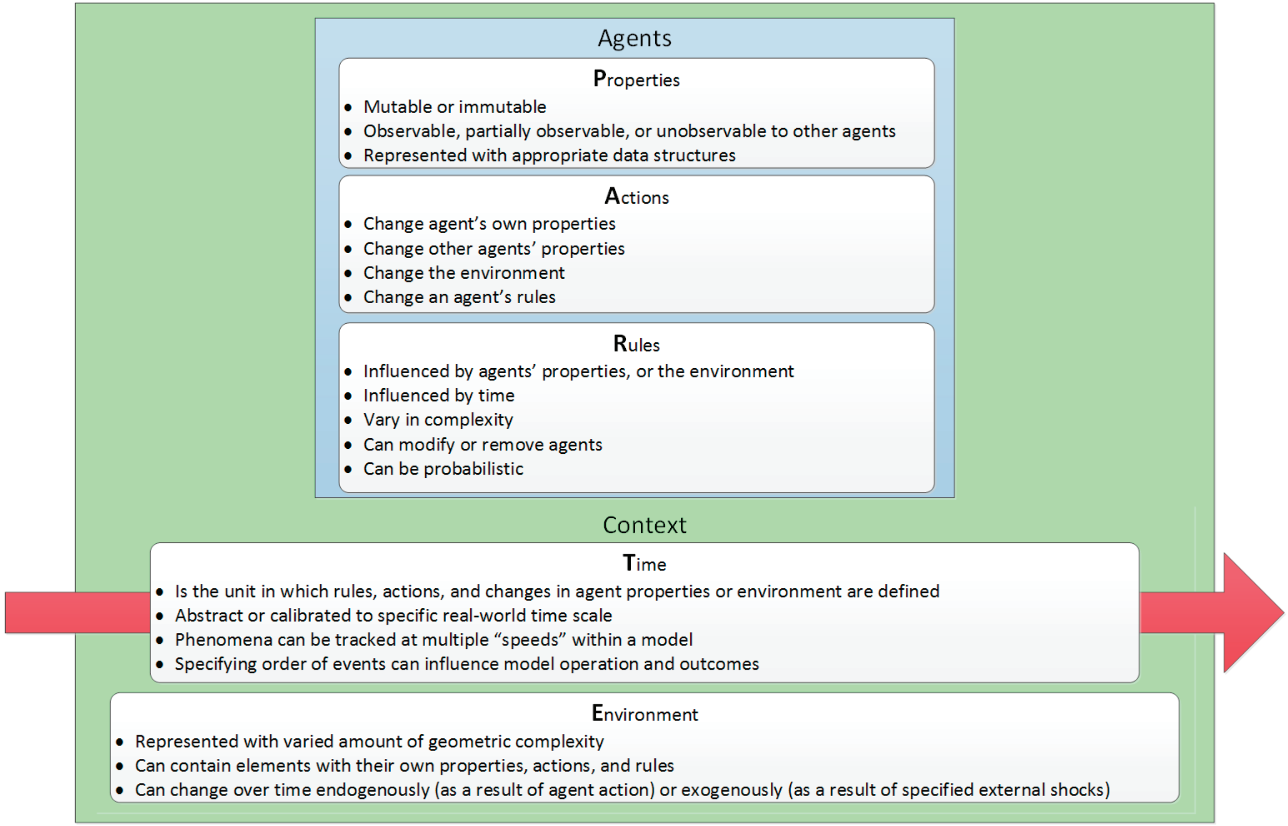

FIGURE A-1 PARTE framework.

simulation might be central to models considering the lifecourse or overlapping generations (Axtell et al., 2002).

- *Observable, partially observable, or unobservable to other agents*. A simulation of farm behavior might treat size in acres as observable to other farms, but income as only partially observable (Brown et al., 2005a; Brown et al., 2005b; Happe et al., 2006; Magliocca et al., 2014; Sun et al., 2014). In a model of tobacco use, agents may be able to observe numerous cues that suggest whether another agent is a current smoker or not, but previous smoking history is likely harder to observe.
- *Stored in a variety of data structures*, from simple booleans to complex lists or arrays. A dichotomous variable can be easily stored in a simple data structure. Storing the mapping between food types and associated reward values (Hammond et al., 2012) or agents' preferences for other types of choices (Kollman et al., 1992, 1997; Maroulis et al., 2014) might require a more complex data structure, such as a hash table or vector.

Not all types of agents represented in a model need to have all properties; for example, the property "market capitalization" is relevant for agents that represent firms but not for agents that represent employees of the firms. All properties must have well-defined conditions for initialization and for change through time. Initialization can involve draws from predefined distributions or from data and may be conditional on values assigned to other properties. By representing each individual actor as a separate software object, ABM allows enormous flexibility to capture heterogeneity across agents in their properties (see section 1.2).

Actions define the repertoire of specific behaviors that agents can perform within the simulation, such as moving around the environment, eating food, smoking tobacco, communicating information to a neighbor, forming a friendship tie, or buying a product. Agent actions can:

- *Change an agent's own Properties*. For example, taking the action "eat" may affect the property "body mass index" over time (Hammond and Ornstein, 2014); taking the action "buy cigarettes" will immediately affect the property "inventory of cigarettes" (Luke et al., 2014).
- *Change the Properties of other agents*. For example, models of cooperation and reciprocity often contain an action "donate (to agent *x*)" which increases a wealth or wellbeing property of agent *x* while decreasing the same property for the actor (Axelrod et al.,

2004; Hammond and Axelrod, 2006a; Hammond and Axelrod, 2006b; Nowak, 2006).

- *Change the Environment*. In models of land use and commons, for example, exploitation of a resource by one agent may reduce the amount of resource available at that location in the environment (temporarily or permanently) to other agents.
- *Change an agent's own Rules*, for example through learning.

For every action included for any agent in an ABM, the modeler must define conditions under which the action is triggered or may be performed. Each action must also have defined consequences (which may include one or more of the changes above); actions that have no consequences do not affect simulation dynamics and do not belong in the model.

Rules are the central drivers of model dynamics, defining how agents choose an action, update properties, and interact with each other and their environment. Rules in an ABM can:

- *Take as an input the current or past value of Properties* (an agent's own, those of others, or those of the environment); for example, "if age > 18, purchase tobacco."
- *Be dependent in some way on Time*, and may involve learning or adaptation—for example, image scoring in reputation models, or the process of preference formation (Hammond et al., 2012; Nowak, 2006).
- *Vary enormously in complexity* from simple heuristics (for example, "when reaching any four-way intersection, always turn left") to detailed internal models or calculations (for example, agents who collect data about the simulated world and optimize over some objective function).
- *Cause not only modification of agents but creation or removal of agents,* for example, creation of offspring in a demographic model or killing of other agents in a model of genocide (Bhavnani and Miodownik, 2009; Epstein, 2002).
- *Involve stochastic probability*, for example, "when reaching any four-way intersection, turn left with probability 50 percent."

Time is central to a dynamic simulation model. Agent-based models (and other related simulation models) generally have a single, lowest-level fundamental unit of time that represents one pass by the computer through the set of instructions that embody the simulation. This is sometimes referred to as an "iteration," a "tick," or a "round." Time in an ABM:

- *Can remain abstract or can be calibrated to real-world time* with additional design work. Such calibration may be easier for some models than for others, depending in part on whether behaviors in the model occur on known time scales. A model in which an agent's age in years changes every 12 iterations or in which agents travel back and forth between home and office every other iteration is relatively easy to calibrate to calendar time (Luke et al., 2014). A model of chronic disease incidence driven by smoking or a model of opinion-change dynamics may require more work to calibrate (Garcia and Jager, 2011).
- *Can involve multiple distinct "speeds" at which change occurs*, for example, the speed with which a virus spreads through social contact versus the speed with which the virus itself evolves.
- *Is the unit in which rules, action, and changes in agent properties or environment are defined.*
- *May also shape the simulation results through decisions about the order in which instructions are followed by the computer*. For example, some types of diffusion models start with a single agent that deviates from the population and calculate the likelihood that the deviation will spread through the population. The choice between two implementations of diffusion—one in which the early adopter influences others before being influenced by its own neighbors and one in which the two directions of influence are evaluated in the reverse order—can result in very different outcomes, such as whether the deviation persists or dies out. A deterministic agent activation order could mistakenly lock the model into one of those outcomes (see section 4.3).

Environment provides the context for agents and their interactions in the model. The flexibility to represent many different types of environment effectively is a major strength of ABM (see section 1.2). The Environment in an ABM can:

- *Range from simple and relatively abstract geometries* (e.g., a lattice, ring, or torus) *to highly complex ones* (often empirically informed) such as a GIS shape file or a network structure.
- *Contain "agent types" itself, with their own properties, actions, and rules*. For example, a model of subsistence agriculture may contain rules for crop regeneration at any particular environmental location that are dependent on farming and harvest practices as well as intrinsic soil and water conditions (see Axtell et al., 2002).

- *Can change over time*, endogenously (as a result of such agent actions as crop rotation) or exogenously (for example, as a result of a policy change or external shock).

The **PARTE** framework describes fundamental building blocks that are present in every ABM, but as the illustrative examples above show, enormous variation is possible in the form that each element (P, A, R, T, and E) takes from model to model. This flexibility is part of the power and potential of ABM, but also underlines the importance of following best practices in making the numerous design choices required (see section 4 below).

4.2 Key Steps in Agent-Based Modeling

Just as ABMs share key building blocks in common (while exhibiting extensive heterogeneity in instantiation), the process of constructing and using an ABM generally follows six key steps. These steps are not unique to ABM—they are shared with many other forms of computational modeling—but several steps raise particular considerations for ABM. Table A-1 briefly describes the key steps.

With ABM in particular, progress from step to step may not be linear, but instead may involve iterative cycles or back-and-forth—especially as models are built up from simple to complex in stages (see section 4.3 BP3 below).

TABLE A-1 Key Steps in Model Development

Key Step	Description
1. Definition of question or goal	Thoroughly consider and spell out the goal that the model will be designed to serve or the specific questions that it will try to answer.
2. Model scope and conceptual design	Identify key concepts, structures, and relationships from the literature and preliminary studies. Determine clearly defined geographic and temporal contexts that are sufficient to achieve research goals.
3. Model specification	Design the model, operationalizing the model "ingredients" identified in the previous step in an implementation-ready way.
4. Model implementation	Translate the specified model into a computationally operational program. Determine initial model parameter values by using estimates from real-world data and engagement with content domain experts.

TABLE A-1 Continued

Key Step	Description
5. Analysis	
a. Testing and calibration	Test the model against real-world data and, if necessary, iteratively calibrate the model design and parameters.
b. Designing experiments and conducting analysis	Create simulated scenarios that use the model to test hypotheses that are central to the research focus, and interpret their results to explore research questions.
c. Sensitivity analysis	Sweep parameter space to identify key leverage points (i.e., parameter values at which small changes in the system can result in drastic changes in outcomes) and to map the set of assumptions and parameter choices that are inputs into the model onto the set of outcomes that it can produce.
6. Synthesis and reporting	Combine findings from experiments and sensitivity analysis and interpret conceptually. Compile statistical analyses and visualizations of results that clearly depict and document research procedures and findings.

4.3 Considerations and Best Practices for *Model Developers*

This section outlines best practices that have emerged to guide design and use of ABMs. Some are specific to policy-aimed modeling; others are general best practices for good ABM (or even for modeling in general) but may have special relevance or importance when the aim is to inform policy. The best practices are organized according to the six steps of modeling (see section 4.2), and discussed in the subsections below.

1. *Definition of Question or Goal*

Models can be put to many different uses and can help to achieve a number of distinct goals (see section 2.1). The full utility that a model will ultimately have cannot always be foreseen, of course, but models tend to be most useful when they are focused and tailored for a specific purpose. This is because different questions or goals are likely to lead to very different decisions about model structure, different design choices, and different data needs. ABM in particular, because of the specificity that it requires, involves many specific design and implementation choices early in the process (see sections 4.1 and 4.2). Thus, a key best practice is

Best Practice (BP) 1: Start with a clear question or goal, and let this drive early modeling decisions.

Having the question or goal clearly in mind guides the initial steps of model development, including inventory of relevant existing literature, available data, and needed team expertise. Agreement on question and goal between the model design team and potential end users of results can be especially important for policy-oriented modeling. A clear statement of question also helps model designers to ensure that the method chosen is appropriate and well suited—ABM may not always be the best choice (for guidance on when to choose ABM, see section 1.2 above and also Axelrod, 1997a; Axelrod, 2004, 2006a; Axelrod and Tesfatsion, 2006; Heckbert et al., 2010).

2. *Model Scope and Conceptual Design*

This step begins with decisions about model scope and how to best represent key conceptual ideas within the model. Appropriate representation of concepts is an important part of the skill of modeling, is aided by clear questions or goals as guideposts, and works differently depending on the modeling method used. For ABM, the best model designs tend to take as a point of departure consideration of key *actors* in the system rather than an emphasis on variables or *factors* (see Macy and Willer, 2002), as is more usual in, for example public health. An ABM-specific best practice is

BP2: Take an "agent" perspective in initial design, identifying key actors in the system that will be the focus of the model.

Initial model design also involves choices about scope and model boundaries. Here, a common tension occurs in the balance between parsimony and breadth. Models often yield the clearest insights when they remain relatively simple—a principle sometimes referred to as Occam's Razor or the KISS principle (Axelrod, 2006a). Parsimony allows effective tracing from inputs via specific mechanisms to outputs of interest (giving clear answers to "why" and "how" key results obtain; see BP11). Keeping the model simple also helps in managing more pragmatic challenges, such as computational speed or tractability. However, especially with complex-system models, there is often pressure to include as much realism as possible; indeed, part of the motivation for selecting an ABM approach is the increased flexibility that it offers to capture realism and interaction.

Managing this tension is a key part of initial model design. Decisions about what to include in the model and what to leave out are guided in part by clarity in the question statement (see BP1), but a key best practice for ABM is

BP3: Start with (relatively) simple models and build up complexity iteratively, one step at a time.

A common experience with ABMs is that simple models are rich enough to generate complex dynamics, counterintuitive surprises, and important insights. And even simple models involve many distinct design and implementation choices (see section 4.1), which require careful sensitivity analysis and testing (see BP9). Of course, what *simple* means may be contextual and depend in part on the starting point in existing studies and the ultimate goal. Even when a more complex model is envisioned, however, starting simple is usually the right choice. By building up complex models from simple ones, one moving piece at a time, the modeler can maintain clarity about how each piece affects results (see BP11) and can greatly facilitate interpretation (see BP12).

When the goal of the modeling effort is *prospective* or *retrospective* policy assessment, an additional design consideration is appropriate representation of policies themselves. Engagement with stakeholders can be an important input into this type of model to ensure, for example, that policies considered in the model are "realistic" ones (e.g., of interest in the real world). At the same time, the modeler must maintain enough independence from stakeholder concerns to avoid building foregone conclusions into the model. (For more on working with stakeholders, see IOM and NRC, 2015.)

3. *Model Specification*

This step involves operationalizing the model "ingredients" in an implementation-ready way, moving from a conceptual design to a specific and explicit sketch of the model. For an ABM, this involves fully specifying P, A, R, T, and E (see section 4.1).

Part of the power of computational and mathematical models comes from a clear and explicit statement of the assumptions that drive results (see section 2.1), and models are only as good as their assumptions. The ability of a model to provide clear and convincing insights thus depends critically on supporting assumptions. An important best practice in model specification is

BP4: Each assumption should be well grounded and have a strong motivation.

Assumptions can be grounded in data, grounded in theory (or external face validity), and (sometimes) included for the specific purpose of considering the sensitivity of model results to their formulation (e.g., as in hypothetical policy scenarios). Regardless of their origin and grounding, full sensitivity analysis is needed for all assumptions (see BP9).

Recourse to a sufficiently interdisciplinary group of content experts can be critical in developing a well-grounded specification and a model that can meet its goals. Pragmatic consideration of opportunities for testing or calibration (including data availability) may also be important at this stage, depending on the stated goals for the model. If the goal includes empirical explanatory power or forecasting, for example, inclusion of model outcomes for which no data are available may be problematic; if the goal is to develop theory or design experiments, this may be desirable instead.

4. *Model Implementation*

This step involves translation of the specific model into an operational form to conduct simulations. For an ABM, this involves writing computer code—either from scratch, or using one of a number of packages that provide some functionality for routine tasks (see Axelrod and Tesfatsion, 2006; Axelrod, 1997b; Rand and Rust, 2011, for discussion of these). Translation from prose descriptions, schematics, or "pseudocode" to formal mathematical and computational instructions is challenging and requires close attention for a number of reasons.

BP5: Use care in translation of an ABM design into computational code.

Computers require very specific instructions and cannot "fill in" any gaps—initial attempts to implement a model often lead to the discovery that the specification (step 3) is insufficiently detailed and requires further thought. Specific choices of functional form or algorithm, required for computational implementation, can affect the results of an ABM and may require additional consideration at this stage (see also BP9). If the person doing the computer coding is not the same person as the model designer in steps 2 and 3, there is a danger of miscommunication or "loss in translation" (see Axelrod, 2006a; Axelrod, 2006b). Additional considerations may arise from the choice of computer language or modeling package. Because software for developing ABM is not standardized and is often open-source and continually evolving, it is important to check whether any design choices or algorithms are "hard-coded" by default. The implementation stage can also lead to tension between the conceptual design and goals of the model (steps 1–3) and the capabilities of the software platform. Although pragmatic considerations concerning feasibility and effort can be real constraints, it is important to avoid letting the available software tools drive the model design away from the imperatives of goal and question.

Given the importance of an implementation that accurately reflects design, and the number of specific choices involved in implementing an ABM, it is an important best practice to conduct several rounds of error-checking and "partial model testing" once code has been written.

BP6: Conduct error-checking and partial testing as models are implemented.

There are many approaches to this type of testing (see Axelrod, 1997a; Axtell et al., 1996; Miller and Page, 2007; Rand and Rust, 2011; for discussion), which has two fundamental goals. The first is to ensure accurate translation from conceptual to computational, and to catch any errors in coding—this often involves both review of the computational code line by line, and the design and application of simple tests of functionality that match actual computational outputs from small pieces of the code with those expected. The second goal of partial testing is to ensure that the model specification itself (assuming proper implementation in code) represents concepts and meets design goals appropriately. Boundary-adequacy tests and extreme-event tests can help to uncover flaws in the model specification that result in dynamics that, for example, violate face validity or clash with conceptual design and require revisiting step 3 in the process. The ability to conduct this type of testing effectively is another important reason to build model complexity slowly (see BP3).

For ABM in particular, there are many details of implementation that can strongly shape dynamics. These may include interaction topology, agent activation regime, randomization of lists, and handling of pseudorandom generation (see Axtell, 2000); each of these topics deserves consideration in the implementation step but may also require sensitivity testing (see BP9). Of particular importance for ABM are decisions about initialization and halting conditions. Every property included in agents will require a starting (initialization) value in the computer, and generating results from a simulation requires instructions to the computer about *when* (in Time) to stop the simulation and calculate the outputs. Both decisions can affect results, and they require special care and consideration but sometimes do not arise until the implementation phase. Like other assumptions, these decisions should be grounded and have a strong motivation (see BP4).

During implementation, documentation of all the specific decisions made becomes a key best practice. Models can go through several iterations of design, specification, and implementation, and maintaining alignment between the actual computer code and the description (in prose or mathematics) is critical, as is version numbering to ensure a match between any particular set of results and the exact code that generated them.

BP7: Fully document model specification and implementation, and maintain up-to-date documentation throughout the process.

Documentation should cover not only the key ingredients in the model specification (P, A, R, T, and E) but also specific choices in implementation, such as those described above. Given widely varying standards across journals and fields about source-code availability, and the variation in packages for ABM, the documentation should aim where possible to be precise enough to allow replicability on its own. Use of an open-source programming language or package, and provision of programming code for published ABMs are also important best practices.

5. *Analysis*

Once the model has been fully implemented computationally, it can be used to conduct analysis. Depending on the goals and questions (see BP1), the analysis may take a variety of different forms. For most of these, an important early step will be testing and/or calibrating the model. There are many approaches to testing (see Epstein, 2012; Heckbert et al., 2010; Manson and Evans, 2007; Miller and Page, 2007; Rand and Rust, 2011), which may involve "stylized facts" from published literature, primary data collection, or use of secondary data, such as those from surveys and experiments, GIS data, and surveillance data. It is important that testing and calibration procedures be consistent with the goal of the model and the question being considered, so it is often important to consider testing from the very outset of design.

BP8: Undertake carefully considered testing and calibration of the model, consistent with the goal or question.

Testing of a model often focuses on comparing outputs with reference data but may also involve comparison or manipulation of inputs (see Rand and Rust, 2011). All procedures and datasets used in testing or calibration should be part of the documentation for the project.

For almost any question or goal, a key part of analysis using an ABM is *sensitivity analysis*. This process involves testing the dependency of model outputs to variation in each of the inputs (assumptions and parameters) and sometimes specific implementation choices.

BP9: Conduct thorough and appropriate sensitivity analysis.

A good sensitivity analysis will usually go beyond testing inputs one by one, instead co-varying inputs over wide ranges to understand sensitivity to differing combinations of parameters. Special attention may be needed both to halting conditions and to initialization (see BP6). Although increasing computational power makes conducting thorough sensitivity analysis easier, the importance of this step can act as a practical limit on model complexity and helps to motivate BP3. The design of (and results from) sensitivity analysis should be well documented and should serve the central goals of internal consistency and increased confidence in the robustness of results being reported.

Once a model has been implemented (and often after testing), it can be put to use. Many models are designed to yield specific insights or answer specific questions. Serving this goal requires designing clear *experiments* to conduct in the artificial world of the model.

BP10: Design clear experiments to yield clear insights.

The accessibility and flexibility of ABM can lead to a temptation to "explore" the model's behavior in an undirected way (Macy and Willer, 2002), but this rarely yields clear insights and can quickly become overwhelming. Thinking carefully about the questions of interest (see BP1) and how to design appropriate experiments to generate clear answers within the model is an important best practice for using ABM effectively and efficiently. For policy-oriented models, this may involve consideration of how to represent a "policy" or "intervention" in the model appropriately. Just as with model implementation, *documentation* of the experiments conducted with the model (specific parameterizations, code version used, and so on) is critical.

Results from model analysis (whether in testing, sensitivity analysis, or experimentation) can sometimes be surprising or counterintuitive. This can occur even in simple models (see BP3) but especially in more complex ones. A critical best practice is to investigate surprises so that why and how they arise can be understood (Axelrod, 2006a,b).

BP11: Always investigate surprising results, and make sure that you understand how they arise.

This may require additional work (and even new elements of code to help track internal states of the model), but is crucial both to ensure that errors are caught and to effectively communicate complex and surprising results by providing intuition to accompany them.

6. *Synthesis and Reporting*

Once a model has been implemented, tested, and analyzed, the next step is to interpret the findings. Drawing appropriate conclusions from simulation results is not always straightforward, especially when models are stochastic, involve numerous inputs, and include multiple mechanisms. Caution is needed to avoid overclaiming (or underclaiming) and to convey appropriate nuance and uncertainty in findings (see BP14).

BP12: Draw appropriate conclusions from the model analysis.

Forthright disclosure of findings, design issues, and sensitivity of results to input assumptions is important. Transparency in (and documentation of) the process used to design, implement, and analyze the model is important. Engagement with subject-matter experts or stakeholders may be needed at the stage of conceptual interpretation.

For almost any goal, an important step for a computational model such as an ABM is to translate the quantitative output of the simulation back into conceptual language that is appropriate for the intended audience. This may involve connecting the model and its results to an existing literature or conversation. ABM in particular can often lend itself to very visual depictions of model dynamics, and designing and executing effective visualization can often be a time-consuming process (and may involve additional computer programming).

BP13: Visualize and translate results into conceptual language.

For policy-oriented use, particular care is needed to ensure that visualizations and conceptual descriptions of model findings are designed with the likely audience in mind and are accurately representing the modeling results (including nuance and uncertainty in findings). Visualization and conceptual description may also cover the analysis and testing procedures used.

Tension can arise with complex models between the goal of descriptive accuracy and the goal of clear communication to a nontechnical audience (Happe et al., 2006). Managing this tension is facilitated by starting with simple models and building up complexity in layers with a clear sense of the contribution of each layer to the outcomes (see BP3 and Macy and Willer, 2002; Tesfatsion and Judd, 2006).

A central issue in conveying models and their results, particularly in a policy context, is managing and communicating uncertainty appropriately. This involves first quantifying uncertainty and its origins—often a distinction is made between "aleatory variability" (natural randomness in a process that cannot be removed) and "epistemic uncertainty" (driven by limited data or knowledge) (see Berger and Troost, 2014). Uncertainty must then be effectively communicated along with the model design and results, and expectations about accuracy in forecasting or policy assessment must be managed.

BP14: Manage and communicate uncertainty appropriately.

Even the best models almost never remove uncertainty and the need for judgment in interpreting results and applying them to real-world situations. Conveying the degree of uncertainty, and its nature and source, is often a key task for a modeler working in a policy context. Recognition (both by modelers and by model consumers) that models are just one input in the decision-making process is important.

The use of modeling to inform a decision process may go beyond the design, execution, and interpretation of any single model. For complex real-world problems and decisions, the use of multiple models or multiple methods can be particularly helpful (see section 2.1 discussion of modeling networks, such as MIDAS).

BP15: Consider multiple models or methods in the context of a broader decision-making process.

Models may be designed independently to answer the same question (giving additional confidence where they agree) or may be designed to complement one another by covering different parts of a topic to preserve parsimony within each individual model while increasing the scope of the overall effort. Models may also sometimes be linked directly (for example, outputs of one model used as inputs in another model), but this requires consideration early in the design process.

4.4 A Few Considerations for *Model Consumers*

The sections above have described the many elements (section 4.1), steps (section 4.2), and best practices (section 4.3) involved in constructing and using ABM. With these in mind, a few guidelines arise for decision-makers who wish to use modeling as an input into the decision process.

Early engagement with the modeling effort can be helpful in communicating the goal or question of interest to model designers and in ensuring that the fit between the desired use of the model and the method and design of the model is appropriate. Model consumers are not always involved in the design and implementation phases of modeling, but they can be. Engagement in design itself can take many forms but often involves helping to ensure "face validity" and relevance of key design choices. Model consumers should ask the right questions throughout the process to ensure that they understand the choices being made.

Later in the process, model consumers should also ask the right questions to ensure that they fully understand the results and their boundaries and possible interpretations, the sensitivity of results to assumptions, and the role of uncertainty. It can be especially helpful for decision makers to understand the intuition and pathways behind results that seem counterintuitive.

4.5 Common Misperceptions About Agent-Based Modeling

A few misperceptions about the use of ABM commonly arise, especially in its use for policy purposes, and are worth brief discussion and clarification. One common misperception is that ABMs are necessarily "ad hoc" or reliant on poorly grounded inputs and assumptions. ABMs certainly *can* suffer from this problem (as can models of all types!) but they *needn't*—there is nothing inherent in the ABM method that prohibits well-grounded assumptions. As described above (see section 4.3), care is needed in model design to motivate and ground assumptions and to avoid growing models too rapidly in complexity and stretching the ability of the modeling team to defend assumptions and explore sensitivity. The flexibility and individual-level focus of ABM confer great power, but they also require careful attention to and responsibility for assumptions on the part of the modeler.

A second common misperception concerns reuse of models. As described above, good models are usually designed for quite specific purposes with clear questions and boundaries in mind, and modelers make many specific implementation choices that flow from these goals. One implication of this is that "repurposing" models to answer questions or address topics and contexts for which they were not designed must be done with great care. Models *can* be used effectively in this way, but it requires carefully revisiting assumptions and design choices to ensure that they remain appropriate for the new application.

A third common misconception is in regard to the skill set required for ABM. Although ABMs are computational models, their rigorous design and use require much more than the skill of computer programming or computer science. Navigating the many elements and challenges of design, implementation, and interpretation of ABM requires another skill, the skill of "modeling," and benefits from extensive experience and familiarity with the best practices outlined above.

5. CONCLUSION

This paper has reviewed the many potential uses of ABM to inform policy or decision making, the features of the technique that make it compelling to use for such purposes, and the best practices involved in doing so responsibly and rigorously. The central message of the paper is that the use of ABM as an input into the policy process is promising and practical, but it is also challenging and complex. As the use of ABM in this way continues to become more widespread, I hope that the overview of key considerations given here will contribute to careful and appropriate use of this powerful tool.

REFERENCES

Axelrod, R. 1997a. Resources for agent-based modeling. In *The complexity of cooperation: Agent-based models of competition and collaboration*. Princeton, NJ: Princeton University Press. Pp. 206–222.

———. 1997b. *The complexity of cooperation: Agent-based models of competition and collaboration*. Princeton, NJ: Princeton University Press.

———. 2004. *Comparing modeling methodologies*. Ann Arbor, MI: Gerald R. Ford School of Public Policy.

———. 2006a. Advancing the art of simulation in the social sciences. In *Handbook of research on nature-inspired computing for economics and management,* edited by J. P. Rennard. Hershey, PA: Idea Group Reference. Pp. 90–110.

———. 2006b. Agent-based modeling as a bridge between disciplines. In *Handbook of computational economics,* edited by L. Tesfatsion and K. L. Judd. Amsterdam, Netherlands: North Holland Publishing Co. Pp. 1565–1584.

Axelrod, R., and L. Tesfatsion. 2006. Guide for newcomers to agent-based modeling in the social sciences. In *Handbook of computational economics,* edited by L. Tesfatsion and K. L. Judd. Amsterdam, Netherlands: North Holland Publishing Co. Pp. 1647–1659.

Axelrod, R., R. A. Hammond, and A. Grafen. 2004. Altruism via kin-selection strategies that rely on arbitrary tags with which they coevolve. *Evolution* 58(8):1833–1838.

Axtell, R. 2000. *Effects of interaction topology and activation regime in several multi-agent systems*. Washington, DC: The Brookings Institution.

Axtell, R. L., and J. M. Epstein. 1999. *Coordination in transient social networks: An agent-based computational model of the timing of retirement*. Washington, DC: The Brookings Institution.

Axtell, R., R. Axelrod, J. M. Epstein, and M. D. Cohen. 1996. Aligning simulation models: A case study and results. *Computational & Mathematical Organization Theory* 1(2):123–141.

Axtell, R. L., J. M. Epstein, J. S. Dean, G. J. Gumerman, A. C. Swedlund, J. Harburger, S. Chakravarty, R. Hammond, J. Parker, and M. Parker. 2002. Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley. *Proceedings of the National Academy of Sciences of the United States of America* 99(Suppl 3):7275–7279.

Bahr, D. B., R. C. Browning, H. R. Wyatt, and J. O. Hill. 2009. Exploiting social networks to mitigate the obesity epidemic. *Obesity (Silver Spring)* 17(4):723–728.

Bendor, J., and T. M. Moe. 1985. An adaptive model of bureaucratic politics. *The American Political Science Review* 79(3):755–774.

Bendor, J., D. Diermeier, and M. Ting. 2003. A behavioral model of turnout. *American Political Science Review* 97(2):261–280.

Berger, T., and C. Troost. 2014. Agent-based modelling of climate adaptation and mitigation options in agriculture. *Journal of Agricultural Economics* 65(2):323–348.

Berger, T., R. Birner, N. McCarthy, J. Díaz, and H. Wittmer. 2007. Capturing the complexity of water uses and water users within a multi-agent framework. *Water Resources Management* 21(1):129–148.

Bhavnani, R., and D. Miodownik. 2009. Ethnic polarization, ethnic salience, and civil war. *Journal of Conflict Resolution* 53(1):30–49.

Brady, M., C. Sahrbacher, K. Kellermann, and K. Happe. 2012. An agent-based approach to modeling impacts of agricultural policy on land use, biodiversity and ecosystem services. *Landscape Ecology* 27(9):1363–1381.

Brown, D. G., S. Page, R. Riolo, M. Zellner, and W. Rand. 2005a. Path dependence and the validation of agent-based spatial models of land use. *International Journal of Geographical Information Science* 19(2):153–174.

Brown, D. G., R. Riolo, D. T. Robinson, M. North, and W. Rand. 2005b. Spatial process and data models: Toward integration of agent-based models and GIS. *Journal of Geographical Systems* 7(1):25–47.

Bruch, E. E., and R. D. Mare. 2006. Neighborhood choice and neighborhood change. *American Journal of Sociology* 112(3):667–709.

Burke, D. S., J. M. Epstein, D. A. Cummings, J. I. Parker, K. C. Cline, R. M. Singa, and S. Chakravarty. 2006. Individual-based computational modeling of smallpox epidemic control strategies. *Academic Emergency Medicine* 13(11):1142–1149.

COSSA (Consortium of Social Science Associations). 2014. *OBSSR conference held on complex systems, health disparities & population health: Building bridges*. http://www.cossa.org/volume33/OBSSRsystems.pdf (accessed December 18, 2014).

Dawid, H. and G. Fagiolo (Eds.). 2008. Agent-based models for economic policy design [Special issue]. *Journal of Economic Behavior & Organization* 67(2).

Dawid, H., and M. Neugart. 2011. Agent-based models for economic policy design. *Eastern Economic Journal* 37(1):44–50.

DeAngelis, D. L., and W. M. Mooij. 2005. Individual-based modeling of ecological and evolutionary processes. *Annual Review of Ecology, Evolution, and Systematics* 36:147–168.

Epstein, J. M. 2002. Modeling civil violence: An agent-based computational approach. *Proceedings of the National Academy of Sciences of the United States of America* 99(Suppl 3):7243–7250.

———. 2004. *Toward a containment strategy for smallpox bioterror: An individual-based computational approach*. Washington, DC: Brookings Institution Press.

———. 2008. Why model? *Journal of Artificial Societies & Social Simulation* 11(4):12.

———. 2009. Modelling to contain pandemics. *Nature* 460(7256):687.

———. 2012. *Generative social science: Studies in agent-based computational modeling*. Princeton, NJ: Princeton University Press.

Epstein, J. M., R. Pankajakshan, and R. A. Hammond. 2011. Combining computational fluid dynamics and agent-based modeling: A new approach to evacuation planning. *PLoS ONE* 6(5):e20139.

Epstein, J. M., J. Parker, D. Cummings, and R. A. Hammond. 2008. Coupled contagion dynamics of fear and disease: Mathematical and computational explorations. *PLoS ONE* 3(12):e3955.

Eubank, S., H. Guclu, V. A. Kumar, M. V. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang. 2004. Modelling disease outbreaks in realistic urban social networks. *Nature* 429(6988):180–184.

Farmer, J. D. 2000. A simple model for the nonequilibrium dynamics and evolution of a financial market. *International Journal of Theoretical and Applied Finance* 3(3):425–441.

Farmer, J. D., and D. Foley. 2009. The economy needs agent-based modelling. *Nature* 460(7256):685–686.

FDA (U.S. Food and Drug Administration). 2013. Modeling and statistical methods for the regulatory assessment of tobacco products—a public workshop, December 6, 2013, Rockville, MD.

Ferguson, N. M., D. A. T. Cummings, C. Fraser, J. C. Cajka, P. C. Cooley, and D. S. Burke. 2006. Strategies for mitigating an influenza pandemic. *Nature* 442(7101):448–452.

Frederick, R. 2013. Agents of influence. *Proceedings of the National Academy of Sciences of the United States of America* 110(10):3703–3705.

Garcia, R., and W. Jager. 2011. From the special issue editors: Agent-based modeling of innovation diffusion. *Journal of Product Innovation Management* 28(2):148–151.

Geanakoplos, J., R. Axtell, J. Farmer, P. Howitt, B. Conlee, J. Goldstein, M. Hendrey, N. Palmer, and C.-Y. Yang. 2012. Getting at systemic risk via an agent-based model of the housing market. *American Economic Review* 102(3):53–58.

Germann, T. C., K. Kadau, I. M. Longini, and C. A. Macken. 2006. Mitigation strategies for pandemic influenza in the United States. *Proceedings of the National Academy of Sciences of the United States of America* 103(15):5935–5940.

Glass, T. A., and M. J. McAtee. 2006. Behavioral science at the crossroads in public health: Extending horizons, envisioning the future. *Social Science and Medicine* 62(7):1650–1671.

Guzy, M. R., C. L. Smith, J. P. Bolte, D. W. Hulse, and S. V. Gregory. 2008. Policy research using agent-based modeling to assess future impacts of urban expansion into farmlands and forests. *Ecology and Society* 13(1):37.

Hall, K. D. 2010. Predicting metabolic adaptation, body weight change, and energy intake in humans. *American Journal of Physiology-Endocrinology and Metabolism* 298(3):E449–E466.

Hall, K. D., R. A. Hammond, and H. Rahmandad. 2014. Dynamic interplay among homeostatic, hedonic, and cognitive feedback circuits regulating body weight. *American Journal of Public Health* 104 (7):1169–1175.

Hammond, R. 2008. Endogenous transition dynamics in corruption: An agent-based computer model. Center on Social and Economic Dynamics. Working Paper No. 19, revised. Washington, DC: Brookings Insitution.

———. 2009. Complex systems modeling for obesity research. *Preventing Chronic Disease* 6(3):A97.

Hammond, R. A., and R. Axelrod. 2006a. Evolution of contingent altruism when cooperation is expensive. *Theoretical Population Biology* 69(3):333–338.

———. 2006b. The evolution of ethnocentrism. *Journal of Conflict Resolution* 50(6):926–936.

Hammond, R. A., and J. T. Ornstein. 2014. A model of social influence on body mass index. *Annals of the New York Academy of Sciences* 1331:34–42.

Hammond, R. A., J. T. Ornstein, L. K. Fellows, L. Dube, R. Levitan, and A. Dagher. 2012. A model of food reward learning with dynamic reward exposure. *Frontiers in Computational Neuroscience* 6(82):1–8.

Happe, K., K. Kellermann, and A. Balmann. 2006. Agent-based analysis of agricultural policies: An illustration of the agricultural policy simulator Agripolis, its adaptation and behavior. *Ecology and Society* 11(1):49.

Happe, K., A. Balmann, K. Kellermann, and C. Sahrbacher. 2008. Does structure matter? The impact of switching the agricultural policy regime on farm structures. *Journal of Economic Behavior & Organization* 67(2):431–444.

Heckbert, S. 2011. Agent-based modelling of emissions trading for coastal landscapes in transition. *Journal of Land Use Science* 6(2-3):137–150.

Heckbert, S., T. Baynes, and A. Reeson. 2010. Agent-based modeling in ecological economics. *Annals of the New York Academy of Sciences* 1185(1):39–53.

Holland, J. 1992. *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control and artificial intelligence*. Cambridge, MA: MIT Press.

Huang, T. T. K., B. Grimm, and R. A. Hammond. 2011. A systems-based typological framework for understanding the sustainability, scalability, and reach of childhood obesity interventions. *Children's Health Care* 40(3):253–266.

IOM (Institute of Medicine). 2010. *Bridging the evidence gap in obesity prevention: A framework to inform decision making*. Edited by S. K. Kumanyika, L. Parker, and L. J. Sim. Washington, DC: The National Academies Press.

———. 2012. *Accelerating progress in obesity prevention: Solving the weight of the nation*. Washington, DC: The National Academies Press.

———. 2013. *Evaluating obesity prevention efforts: A plan for measuring progress*. Washington, DC: The National Academies Press.

IOM and NRC (National Research Council). 2015. *A framework for assessing effects of the food system*. Edited by M. C. Nesheim, M. Oria, and P. T. Yih. Washington, DC: The National Academies Press.

Kollman, K., J. H. Miller, and S. E. Page. 1992. Adaptive parties in spatial elections. *American Political Science Review* 86(4):929–937.

———. 1997. Political institutions and sorting in a Tiebout model. *The American Economic Review* 87:977–992.

Laver, M. 2005. Policy and the dynamics of political competition. *American Political Science Review* 99(2):263–281.

Laver, M., and E. Sergenti. 2011. *Party competition: An agent-based model*. Princeton, NJ: Princeton University Press.

LeBaron, B. and P. Winkler (Eds.). 2008. Agent-based models for economic policy advice [Special issue]. *Journal of Economics and Statistics* 228(2–3).

Lee, B. Y., S. T. Brown, G. Korch, P. C. Cooley, R. K. Zimmerman, W. D. Wheaton, S. M. Zimmer, J. J. Grefenstette, R. R. Bailey, T. M. Assi, and D. S. Burke. 2010. A computer simulation of vaccine prioritization, allocation, and rationing during the 2009 H1N1 influenza pandemic. *Vaccine* 28(31):4875–4879.

Longini, I. M., A. Nizam, S. Xu, K. Ungchusak, W. Hanshaoworakul, D. A. T. Cummings, and M. E. Halloran. 2005. Containing pandemic influenza at the source. *Science* 309(5737):1083–1087.

Longini Jr, I. M., M. Elizabeth Halloran, A. Nizam, Y. Yang, S. Xu, D. S. Burke, D. A. Cummings, and J. M. Epstein. 2007. Containing a large bioterrorist smallpox attack: A computer simulation approach. *International Journal of Infectious Diseases* 11(2):98–108.

Luke, D. A., A. A. Sorg, A. Mack-Crane, R. A. Hammond, M. E. Kasman, K. M. Ribisl, and L. Henriksen. 2014. "Tobacco Town: Modeling the effects of tobacco retail reduction." Poster presentation at Annual Investigators Meeting, State and Community Tobacco Control Initiative, National Cancer Institute, Chicago, IL.

Mabry, P. L., S. E. Marcus, P. I. Clark, S. J. Leischow, and D. Méndez. 2010. Systems science: A revolution in public health policy research. *American Journal of Public Health* 100(7):1161–1163.

Macy, M. W., and R. Willer. 2002. From factors to actors: Computational sociology and agent-based modeling. *Annual Review of Sociology* 28(1):143–166.

Magliocca, N. R., D. G. Brown, and E. C. Ellis. 2014. Cross-site comparison of land-use decision-making and its consequences across land systems with a generalized agent-based model. *PLoS ONE* 9(1):e86179.

Manson, S. M., and T. Evans. 2007. Agent-based modeling of deforestation in southern Yucatán, Mexico, and reforestation in the midwest United States. *Proceedings of the National Academy of Sciences of the United States of America* 104(52):20678–20683.

Maroulis, S., E. Bakshy, L. Gomez, and U. Wilensky. 2014. Modeling the transition to public school choice. *Journal of Artificial Societies and Social Simulation* 17(2):3.

Miller, J. H., and S. E. Page. 2007. *Complex adaptive systems: An introduction to computational models of social life*. Princeton, NJ: Princeton University Press.

Nader, P. R., T. T. K. Huang, S. Gahagan, S. Kumanyika, R. A. Hammond, and K. K. Christoffel. 2012. Next steps in obesity prevention: Altering early life systems to support healthy parents, infants, and toddlers. *Childhood Obesity (Formerly Obesity and Weight Management)* 8(3):195–204.

North, M. J., C. M. Macal, J. S. Aubin, P. Thimmapuram, M. Bragen, J. Hahn, J. Karr, N. Brigham, M. E. Lacy, and D. Hampton. 2010. Multiscale agent-based consumer market modeling. *Complexity* 15(5):37–47.

Nowak, M. A. 2006. Five rules for the evolution of cooperation. *Science* 314(5805):1560–1563.

Ohtsuki, H., C. Hauert, E. Lieberman, and M. A. Nowak. 2006. A simple rule for the evolution of cooperation on graphs and social networks. *Nature* 441(7092):502–505.

Page, S. E. 1999. Computational models from A to Z. *Complexity* 5(1):35–41.

———. 2006. Path dependence. *Quarterly Journal of Political Science* 1(1):87–115.

Rand, W., and R. T. Rust. 2011. Agent-based modeling in marketing: Guidelines for rigor. *International Journal of Research in Marketing* 28(3):181–193.

Schelling, T. C. 1971. Dynamic models of segregation. *Journal of Mathematical Sociology* 1(2):143–186.

Schlüter, M., and C. Pahl-Wostl. 2007. Mechanisms of resilience in common-pool resource management systems: An agent-based model of water use in a river basin. *Ecology and Society* 12(2):4.

Sterman, J. D. 2006. Learning from evidence in a complex world. *American Journal of Public Health* 96(3):505–514.

Sun, S., D. C. Parker, Q. Huang, T. Filatova, D. T. Robinson, R. L. Riolo, M. Hutchins, and D. G. Brown. 2014. Market impacts on land-use change: An agent-based experiment. *Annals of the Association of American Geographers* 104(3):460–484.

Tesfatsion, L., and K. L. Judd. 2006. *Handbook of computational economics: Agent-based computational economics*. Vol. 2, Amsterdam, Netherlands: North Holland Publishing Co.

Xie, Y., and X. Zhou. 2012. Modeling individual-level heterogeneity in racial residential segregation. *Proceedings of the National Academy of Sciences of the United States of America* 109(29):11646–11651.

Yang, Y., J. D. Sugimoto, M. E. Halloran, N. E. Basta, D. L. Chao, L. Matrajt, G. Potter, E. Kenah, and I. M. Longini. 2009. The transmissibility and control of pandemic influenza a (H1N1) virus. *Science* 326(5953):729–733.

Zhang, J., L. Tong, P. Lamberson, R. A Durazo, A. Luke, and D. A. Shoham. 2014. Leveraging social influence to address overweight and obesity using agent-based models: The role of adolescent social networks. *Social Science and Medicine* 125:203–213.

Appendix B

Agent-Based Models for Policy Analysis

Lawrence Blume[1]

WHAT IS AN AGENT-BASED MODEL?

An agent-based model (ABM) is a computational simulation model of a many-agent system that captures the behaviors of the system's autonomous agents and their interactions with each other. An ABM is a computational instantiation of a complex adaptive system (CAS). A CAS is a dynamic model that represents individual agents and their collective behavior.[2] In social science applications, agents are usually people. CASs, however, have applications in many different systems, in which agency has many different interpretations. Generally speaking, an agent is a persistent entity that is described by states and behaviors, which are consequences of the agent's state. The agent's state is modified by its interactions with other agents. The agent population is usually not modeled as a "gas" of randomly interacting particles.[3] Instead there will typically be some structure to agents' interactions: The set of others with whom any agent can interact is circumscribed. This structure is usually described by a social network; agents interact only with their neighbors.

---

[1] Cornell University, IHS Vienna, and Santa Fe Institute.

[2] The study of CASs, both theoretically and through computer simulations, was central to the research program of the Santa Fe Institute in the 1980s and 1990s and was stimulated, in particular, by John Holland's work on genetic algorithms and classifier systems. His 1975 book (Holland, 1975) was certainly influential at SFI.

[3] In models describing social phenomena at higher levels of aggregation, the "gas" assumption is common. See the discussion of the SIS model below.

The states of individual agents may behave erratically. Nonetheless, some aggregates of agent behavior may exhibit stable regularities. These regularities are referred to as *emergent behaviors*. The equations describing a CAS are descriptions of behavioral rules for the autonomous agents and descriptions of how they interact. The driving equations of the system are typically at the level of the individual agent, and describe action at temporal scales appropriate to the agent. An emergent property is a regularity in the output of the CAS that appears robustly on a temporal or spatial scale different from those of the driving equations. Emergent properties, the behaviors of the whole, are the objects of interest in a CAS.

An ABM is a computer program that implements a CAS by simulating its behavior. The CAS describes a probability distribution on outcomes for every vector of inputs x and equation parameters p, and the ABM simulates the probability distribution. Each run of the program yields a random draw from the CAS's outcome distribution, and so the empirical distribution of many draws approximates the CAS's outcome distribution.

At the risk of being either too mathematical or too redundant, I will finish this section by recasting a familiar epidemiologic model as a CAS and compare the CAS representation with more familiar representations. The annex contains a more formal mathematical description of CASs.

The SIS model is a textbook model of the spread of a disease. At any moment, the individuals are of two types, or states: susceptible (S) and infected (I). The infected individuals can transmit the disease to the susceptibles. The numbers of people of each type at time t in a fixed population of size N are denoted $S(t)$ and $I(t)$, respectively. The point of the model is to track the path of the population through these states over time. A continuous-time SIS model might presume that the population contains a continuum of agents of mass 1. The population aggregates evolve according to the following differential equation system:

$$\begin{aligned} \frac{dS}{dt} &= -\beta IS + \gamma I \\ \frac{dI}{dt} &= \beta IS - \gamma I \end{aligned} \tag{1}$$

The SIS model is an aggregate-level model of an epidemic with the happy feature that no one dies from the disease. It could also serve as a model of the spread of a rumor (although the well-known SIR epidemiological model would be a better metaphor). The parameter β is the transmission rate, and γ is the recovery rate. A solution to the deterministic model is a function that describes the evolution of $S(t)$ and, by implication, $I(t)$ through time. The SIS model is a "gas model" in that key to its derivation is the assumption that the population contains many individuals and in any unit of time each individual is equally likely to interact with any other individual.

Equation system 1 is meant to model the aggregate behavior of a gas model of individuals—individuals bump into each other randomly. If a susceptible individual and an infected individual collide, the susceptible individual becomes infected with probability β. Furthermore, an infected individual becomes cured with probability γ. The hope is that the differential equations provide a good approximation of a large-population version of the gas model.

The aggregate behavior of the stochastic gas model is often described as a birth–death process. This is a Markov process on the number i of infectives. The time interval h of a single period is so small that at most a single transition takes place in each time interval. That is, if $I(t) = i$ in period t, then $I(t + 1)$ can have only the values $i + 1$ (a "birth," or infection), i, or i 1 (a "death," or recovery). The probabilities of births and deaths when $I(t) = i$ are denoted by p_i and q_i, respectively, and they have the values

$$\begin{aligned} p_i &= \beta \frac{i(N-i)}{N} h \\ q_i &= \gamma i h \end{aligned} \tag{2}$$

The equations (2) describe a stochastic process, a joint probability distribution of the collection of random variables $\{S(t)\}_{t=1}^{\infty}$. A single draw from that distribution is a sample path of the number of susceptibles. A textbook theorem says that the solution to the differential equation approximates the path of the stochastic process uniformly well over any finite time horizon if h is small enough and N large enough. That justifies the use of the differential equation, but only up to a degree, because the asymptotic behavior of the differential equation does not in general approximate the asymptotic behavior of the stochastic process.

A CAS model of the same process describes the circumstance of each individual in the population. Thus, a state, or configuration, of the CAS is a vector of length N in which the nth component describes the state of individual n, S (susceptible) or I (infected). The CAS has a list of rules that describe how each individual responds to interactions with others and to exogenous random events. For the illustrative SIS CAS, all individuals have the same rule, which is described in Figure B-1. On every date, one of three things can happen to a susceptible individual: they can be matched with a healthy individual, event h; with an infected individual, event i; or with no one, event 0. The same three events can happen to an infected individual,

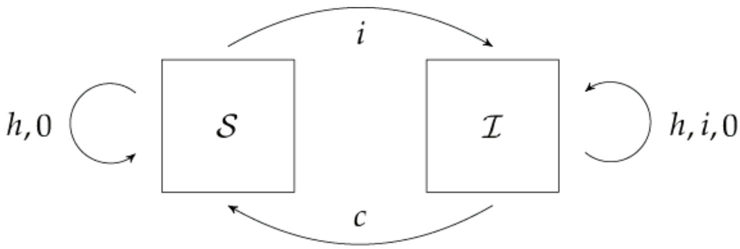

FIGURE B-1 Individual state transitions.

and in addition he or she can be cured, event *c*. The boxes represent their possible states, susceptible, *S*, and infected, *I*. The arcs represent the transitions that each event causes: For instance, if a susceptible individual is matched to another susceptible individual, their state remains unchanged.

To complete the description of the CAS, the interaction process and the exogenous event process must be described. There are $N(N-1)/2$ possible unordered pairs that can form, and there are *I* possible cures. There is also the possibility that nothing will happen. Thus, in any state of the system that has *I* infecteds, there are $N(N-1)/2 + I + 1$ possible events. In each period, taken to be of very short duration *h*, one and only one of those events will be drawn. Suppose that the probability that any particular pair will form is $\beta h/N$, that a cure for a particular infected individual has probability $\gamma h/N$, and that nothing will happen with the complementary probability. Because *h* is small, the pair formation and recovery numbers and 1 minus their sum are nonnegative, so they describe a probability distribution on the set that consists of the pairing events, the recovery events, and the event that nothing happens. Then the stochastic process of the number of susceptibles in the CAS is exactly the birth–death process described by equation 2.

Emergent properties of the CAS have to do with the behavior of aggregates, such as the number of susceptibles. This CAS is a Markov process that has a single absorbing state in which no one is infected; that is, the disease has died out. The distribution of the extinction time, which describes the behavior of the amount of time it takes to reach that state, is another emergent property.

An ABM would implement this CAS in a computer program by iterating the following scheme: Starting with an initial configuration, use a random-number generator to choose a feasible event according to the probabilities described above. If the event is a match, every individual other than the matched pair receives input 0, and each individual in the pair receives the state of their partner. If the event is a cure, the input to the cured individual is the event *c*, and all other individuals receive a 0. The state of each individual is then updated according to the rule of Figure B-1, and the result is a new configuration.

If we are interested only in the aggregates, there is no advantage to constructing the CAS; there are easier ways to simulate the SIS process than by building an ABM. But now make the model more complex. Suppose that individuals are differently susceptible to the disease; that is, β is now individual-specific. Aggregate behavior can still be modeled with a Markov process, in this case a multitype birth–death process, with one type for each level of susceptibility. But such processes are more difficult to analyze, and if every individual has a different susceptibility to the disease, the resulting multitype birth–death process in a population of size N is essentially the CAS. Even more interesting is to suppose that the population has a spatial structure and individuals either meet only neighbors or meet neighbors more frequently than others. Now one needs the CAS to keep track of things. The network becomes a parameter of the CAS, and with an ABM one can ask how, for given β and γ (and h), the shape of the network matters. The CAS can be still more complex. A given set of individuals could be designated as "health care workers" who have higher probabilities of interacting with infected individuals, and so on.

The aggregate behavior of the simple CAS can be usefully approximated for large N and small h over finite time horizons by the differential equation system 1, and a similar system approximates the multitype version when there are many more people than types. If the social structure of the networked model is something regular, like a lattice, it is possible to approximate the system with a partial differential equation if the large-N question is posed the right way. Such approximations are known as mean-field approximations. For more realistic social networks, it is not clear how to pose the large-N question.

CASs force a bottom-up approach to modeling systems. The modeling exercise requires a description of the set of individuals, their behavioral rules, and a description of how they interact. That is in contrast with top-down descriptions, such as equation system 1. One virtue of bottom-up modeling is that the derived aggregate system is guaranteed to be consistent with some actual social process. A further advantage, as this example illustrates, is that bottom-up models support a degree of complexity that is not available in aggregate models. In particular, heterogeneity in agent behaviors and heterogeneity in the variety of interactions available to agents in different roles need more complex descriptions than top-down models can provide.

THE USES OF AGENT-BASED MODELS

ABMs, like other mathematical models, serve three purposes: demonstration, description, and prediction. In the 1980s and 1990s, the primary use of ABMs was to "show off" the kinds of emergent behavior that a system could produce. A good example of that is Schelling's segregation model (Schelling, 1971). It is described by a small number of parameters whose purpose is to show that an emergent property—completely segregated neighborhoods—is a consequence of individual decision rules that exhibit a very small "taste" for similar neighbors. Schelling describes his 1971 paper as "an abstract study of the interactive dynamics of discriminatory individual choice" (p. 143). Many demonstrations are just theoretical exercises in models that cannot (yet) be accessed analytically. I put *yet* in parentheses because analytic technique does advance. Schelling's model could not have been addressed with tools that were available in the late 1960s, but developments in so-called particle systems have made the formal analysis of his model straightforward.[4] Generally speaking, models like Schelling's generate useful theoretical conjectures that can be explored with a variety of methods.

The Hoffer et al. (2009) model of a local heroin street market in Denver is a more sophisticated use of an ABM. The authors use ethnographic data collected by the principal author, an anthropologist, to calibrate an ABM. The ethnographic research concentrated on individual behaviors and interactions of the different actors in the market. The authors make it clear that "the model described in this manuscript is not intended as a forecasting tool" (p. 273). The purpose of the ABM was to uncover emergent properties of market behavior that could not be observed at the scale of ethnographic research. As far as I can tell (the paper is not entirely clear on its methods), the model is calibrated to ethnographic data. Simulations are then run at a market scale to observe the emergent properties of the market system. The model is complex. It contains six types of agents, each with its own rules of interaction and repertoire of behaviors: customers, brokers, sellers, private dealers, police, and homeless people. The customer agent in particular has complex demand behavior that reflects known facts about heroin use. The large-scale properties of the heroin market are likely to be measurable only with great difficulty, so using a simulation model that is based on more easily observed individual behavior patterns is a clever idea. One's confidence in the model's macro-level predictions depends on one's confidence in the internal validity of the model, that is, how well it matches the ethnographic data and how well the ethnographic data capture the fine details of the agent interactions on which the model most sensitively depends. Ideal research design would require dialogue between ethnographers and modelers about the research strategy and the modeling activity. That was impossible in this case because the modeling took place

[4] See Young (2001).

nearly 15 years after the ethnographic research was carried out.[5] The ABM generated some interesting hypotheses about the large-scale behavior of the street market. One interesting experiment is the simulation of a one-night police crackdown. The model showed a short-term effect on all market agents and a sharp decline in transactions. But sales through other channels, particularly private dealers, led to a rebound in transaction volume, and the simulation showed that the police crackdowns had no long-term effect. This detailed model, tightly coupled to a particular market, suggests interesting hypotheses about the effects of policing strategies. The authors are circumspect about the generalizability of their analysis, however, noting both data lapses and peculiarities of the particular market that they studied.

A more ambitious research program is reported in Eubank et al. (2004) and Toroczkai and Eubank (2005). The subtitle of the work by Toroczkai and Eubank (2005) describes a policy problem: "How to halt a smallpox epidemic." An ABM of the spread of smallpox through a city is calibrated on pre-existing data, and then the effects of several different vaccination regimens are simulated. No U.S. city has experienced a smallpox epidemic in recent times, so the data on which the model is calibrated come from a world in which smallpox is absent. The model is then used to simulate counterfactual worlds in which smallpox is spreading (a consequence, presumably, of some biowarfare or terrorist event) and different vaccination strategies are deployed. The model has three components: an urban transportation component, a detailed epidemiological model of smallpox transmission, and a model of disease detection.

The urban transportation simulation model is used to simulate the daily movements of individuals across locations. From this, contacts between individuals are captured and a contact graph is constructed. The urban transportation component is quite complex. It is an ABM of urban transportation designed to describe Portland, Oregon. A synthetic population is constructed whose distribution of demographic and other characteristics matches that observed in Portland census data. Survey data are used to construct activity patterns for households. Activity location is estimated from land-use and transportation-cost data. From this, routes and travel times for each individual are forecast. From this modeling exercise it can be determined that individuals *i* and *j* are in the same location at such-and-such a time. In this manner, a representative contact graph is constructed—who met with whom on a representative day.

The model of smallpox spread from an initial population of infectives is deployed on the contact graph. This too is quite complex, displaying a great deal of biological specificity that I will not describe. The third com-

[5] This comment is not meant as a criticism of an exploratory methodology paper whose purpose was to demonstrate the utility of modeling tools in the ethnographic community.

ponent links actual (simulated) disease spread with the observations that drive these policies.

Four modeling exercises are carried out: A baseline simulation with no treatment, mass vaccination of 100 percent of the population over four days, targeted vaccination and quarantine with "unlimited resources," and a limited resource targeted vaccination and quarantine policy. The conclusion is that targeted response can be effective if detection is sufficiently fast. Mass vaccination is not necessary.

The qualitative result, that policies alternative to mass vaccination could conceivably work, is perhaps interesting. The quantitative results of the simulation exercise, and the conclusion that a targeted vaccination scheme would be as effective as mass vaccination in the Portland of today, are not reliable due to assumptions implicit in the way the model is constructed. For example, we might imagine that knowledge of a spread of smallpox cases would cause people to alter their daily routines. The "representative contact graph" derived from the urban transportation model calibrated to data from a smallpox-free Portland might look very different from a graph of social contacts in a Portland where smallpox is rampant. I will argue below that ABMs that are complex enough to demonstrate or describe possible policy effects in interesting environments will almost certainly fail to measure causal effects to the satisfaction of at least some social science communities.

CAUSATION AND STRUCTURAL MODELS

Empiricists today nearly universally accept Hume's idea that necessary—that is, causal—connections cannot themselves be perceived, that only recurring associations can be observed, and that the fundamental problem of empirical science is to distinguish the causal relationships among all the associations that appear in data. Hume offered two definitions of cause, the second of which has been influential in statistics and the sciences: "We may define a cause to be *an object followed by another, . . . where, if the first object had not been, the second never had existed*."[6] The contemporary instantiations of that idea are counterfactual theories of causation.[7] Those theories consider a number of possible or hypothetical worlds. In a world in which $X = x$ and $Y = y$, the claim that X causes Y is considered by examining nearby worlds in which $X \neq x$. The claim is established if, in worlds that differ only in the assignment of the X value, $Y \neq y$.

[6] Hume (1777) sec. 7, part 2. Italics in the original.

[7] Proponents of this view include the late David Lewis (1973) and Nancy Cartwright (1979, 1989). A recent expression of this theory is Judea Pearl's book (2009). A formal description of some of Pearl's ideas has been developed in Halpern and Pearl (2005a,b).

That loose description of possible-worlds semantics is made rigorous through the use of *structural models*. For our purposes, a structural model contains a system of equations that describes a set of hypotheticals or possible worlds. The possible worlds are ones that satisfy some rules that are meant to be descriptive of the phenomena that the model addresses. For example, in a social-science model, these would include assumptions about how individuals interact and rules for their behavior. The SIS CAS is a structural model. Its equations describe how individuals meet and transmit a disease and how they are cured—a combination of social and biologic rules. The equation system of a structural model contains functional forms, variables, and parameters. Each specification of parameters defines a structure, a possible world, and a specific set of relationships among the variables. The variables themselves are partitioned into exogenous and endogenous variables: those determined outside the model and those determined within the model.[8]

The set of possible worlds to consider in adjudicating causal claims is described by the model. Hume's counterfactual definition of causation leads to a deep point about the nature of causal claims:

> The proposition that it is possible to discover associations among events that are, in fact, invariable ceases to be a provable statement about the natural world and becomes instead a working rule to guide the activity of the scientist. . . . The only "necessary" relationships among variables are the relationships of logical necessity that hold in the scientist's model of the world. . . .
>
> Simon (1953, pp. 49–50)

> Causality is a property of a model. . . .
>
> Heckman (2000, p. 89)

Without a theory, there can be no causal claims. And if all causal claims are relative to particular theories, we are freed from the obligation to find the one true model, the root causes, and can instead look for models that

[8]Equation systems are somewhat arbitrary. An equation $y = ax + u$ can be rewritten as $x = \beta y + \upsilon$ where $\beta = 1/\alpha$ and $\upsilon = u/\beta$. But the causal arrow in the first equation points in the opposite direction from the arrow in the second. The geneticist Sewall Wright (1921, 1925) and later the economist Jan Tinbergen (1968) supplemented their equation systems with a diagram, later called a "path diagram," a graph with nodes that represented variables and directed edges pointing in the direction of causal effects. In his work on causality, Pearl (2009) takes the diagram to be the primitive causal model. Another approach to the problem was taken by the researchers at the Cowles Commission for Research in Economics for the case of linear-equation systems (Hood and Koopmans, 1953). They used transformations of equation systems to determine classes of equivalent systems, all members of which expressed the same causal relationships.

are consistent with the data and address just those questions that we want to ask.

Policy analysis is concerned with three kinds of questions. The first kind asks for the effect of policy X on outcome Y in a given environment E; this is the classic treatment-effects problem. The second is to infer from the effect of X on Y in E the effect of X on Y in a different environment E'. The third is to infer from the effect of X on Y in E the effect of a different policy X' on outcome Y (or even a different outcome Y') in a different environment E'. The last two questions, requiring extrapolation, are fundamentally different from the first in that they require us to use the laws uncovered in the analysis of X and Y in environment E to make predictions about a different environment and perhaps a different policy. For example, in the SIS CAS model, the parameter β is the product of two parameters: β_1, the probability that two individuals meet, and β_2, the probability that the disease is transmitted at a meeting. Suppose that β_1 is a policy variable that can be controlled by, say, a policy of identifying and isolating infected individuals. By estimating β_2 and γ, we can estimate the distribution of the time to extinction of the disease for a given policy β_1. Suppose now that we expect a new variant of the infective agent to sweep through the population with a lower recovery rate γ'. With knowledge of the structural parameter β_2 learned in the initial environment, we can forecast the extinction-time distribution for other policies β'_2 in the new environment γ'.

Unlike the treatment-effects question, these questions require uncovering the parameters of the model, that is, uncovering behavioral laws that remain valid in environments other than E, and in particular environment E', that we want to study. The more "fundamental" or "deeper" the behavior relationships in the model, the larger its domain of applicability, that is, the richer the sets of environments and policies that it can address. For the purposes of policy analysis, the analyst considers a set of possible policies and environments. The model is required to be sufficiently "fundamental" that its relationships are valid for all policies and environments under consideration. I will refer to that as Marschak's stability requirement because Jacob Marschak (1974) was the first to pose it explicitly in a discussion of what makes a good structural model. In contemporary macroeconomics, the phrase *Lucas critique*, in honor of Robert Lucas (1976), is often applied to claims that the behavioral equations of some models are not invariant under alternative macroeconomic policies, that is, that they fail Marschak's stability requirement.

ABMs are structural models. They describe agents' rules for processing information and making choices and for how agents interact. They are useful for demonstrating potential policy outcomes; in particular, they may alert us to emergent consequences of policies that those who design policy, thinking at the behavioral level, may miss. They also allow us to test the reasonableness—the plausibility—of our behavioral assumptions. Nonetheless, they have only limited uses in the statistical analysis of causation and for the extrapolative exercise required for counterfactual policy analysis.[9]

STATISTICAL ISSUES RELATED TO AGENT-BASED MODELS

The purpose of ABMs is to simulate the microbehavior and macrobehavior of CASs, particularly those whose descriptions are too complex to be studied with analytic methods. The necessary complexity of these models makes them large: the smallpox-epidemic ABM (Eubank et al., 2004; Toroczkai and Eubank, 2005), for instance, has over 30 parameters. Those parameters describe different parts of the model, including the course of the disease in a given host (for several variants of the disease), the transmission model (including shedding of an individual's viral load to the environment and uptake by an individual from the environment), the effects of vaccination, and the traffic-simulation tool, which is used to provide a fine-grained description of how the population mixes over a period of weeks.[10]

Using this model to simulate the effects of different vaccination policies requires knowledge of all the parameters. ABM practitioners use such terms as *validation* and *calibration* (inconsistently) to describe methods for choosing parameter values. Tesfatsion (2015) describes three approaches to validation. *Input validation* attempts to use parameter values that come from external knowledge of microbehavior, such as information about the course of smallpox in a single individual; this is the principal validation tool for the disease parts of the smallpox ABM because we have no episodes of smallpox epidemics in recent U.S. history. *Descriptive output validation* matches computationally generated output with preexisting data on the process being modeled; for example, one might fit the SIS ABM to data on chickenpox by choosing parameter values to match data on the time path of chickenpox incidence in a given location and year. Finally, *predictive output validation* matches model outputs to subsequently observed datasets. It

[9] Approaches to causal inference vary across the disciplines in the social sciences. The approaches can be divided into two groups. One group believes causal claims cannot be established without some kind of counterfactual analysis. A more permissive group is comfortable with causal inference from observational data without an implicit or explicit experiment in the background. I stand with the first group. This group itself is divided into those who require a well-motivated model to make causal inferences and those who are willing to infer causation from randomized trials or so-called natural experiments. Although both groups have a lot to argue about, they would certainly agree that is hard to imagine the natural experiment that would identify the effects of alternative vaccination strategies on a smallpox outbreak in a moderate-sized American city. Put more technically, the stable unit treatment value assumption is unlikely to be met for treatment effects one would want to analyze with an ABM.

[10] More recent information about subsequent development of the epidemic simulation modeling tool can be found at http://www.lanl.gov/programs/nisac/episims.shtml.

seems that, generally speaking, two general procedures are used to choose parameter values: choosing them from preexisting studies and choosing them to match statistics of actual and simulated datasets.

The smallpox ABM illustrates a problem that I believe is common in computational models that have many parameters and which makes use of "input validation," that is, choosing parameter values on the basis of out-of-simulation considerations. One would not expect many of the parameters affecting people's travel behavior and contacts to be invariant to the onset of a smallpox outbreak. The failure of parameter invariance to counterfactual initial conditions under consideration, and through a typical run of the model, means that the behavioral relations driving the model are not stable under the counterfactual scenarios that we want the model to examine, so the ABM fails the Marschak criterion. Input validation alone is reasonable for demonstration purposes but not for proof of concept.[11] The problem of parameter stability is often discussed in the context of estimating structural models, but it is even more critical for models that have externally validated, "input-validated" parameters.

Before turning to issues of ABM parameter estimation, I want to mention a fundamental question about the choice of parameter values for policy evaluation: What does it mean to have good parameter estimates? The purpose of ABMs is to study emergent behavior of a system. An empirically successful ABM will get the microbehaviors right, so the agents in the model approximate in some useful way the behaviors of agents in the world. The ABM will also accurately describe the macrobehaviors, the emergent properties. That is an enormous undertaking for any large-scale computational model. A similar problem arises in the currently popular Dynamic Stochastic General Equilibrium (DSGE) computational macroeconomic models. Instead of having many agents, these models attempt to capture individual behaviors with a small number of representative types. Their goal is nonetheless to capture emergent behavior, in this case that of macroeconomic time-series of interest, such as gross domestic product, inflation, and unemployment. DSGE practitioners use the same calibration and validation techniques that are used in ABMs. Early DSGE practitioners uncovered a dilemma: If they calibrated the parameters of the representative agents to values found to be reasonable in microeconomic studies, they would incorrectly forecast the macroeconomic data. On the other hand, calibrating to the macroeconomic data required microeconomically implau-

[11] I chose this model because the subject matter is a typical ABM application and exhibits the complexity that one often sees in ABMs and because publication in *Nature* and evidence of successful grant applications suggest that it is not regarded as a horrible exemplar of an ABM put to the purpose of counterfactual analysis. I did not cherry-pick the model, and I do not believe it to be different in kind from many other ABMs that have been studied.

sible parameter values. Validation poses a tension between different ways of "getting it right." In principle, one would get right both the microbehavior and the macro-level behavior, but this may be impossible. If so, the modeler has to make a choice. DSGE macroeconomists chose to sacrifice behavioral realism to make the macro-level behavior most closely match the data.[12] But doing so makes the whole exercise nothing more than elaborate curve-fitting. Although ABMs may be satisfactory for demonstration purposes, this problem makes them bad policy analysis tools, for two reasons: First, many things that we might wish to calculate, such as the agents' economic welfare or utility, depend on the microparameters, and calculations with the wrong microparameters are likely not to remain stable in counterfactual scenarios—Marschak's stability requirement again.

IDENTIFICATION

It would seem, putting aside the calibration and validation problems raised in the preceding paragraph, that ABMs are useful for predicting the effects of novel policies in complex environments. Ironically, however, the virtues of ABMs—their expressiveness, their ability to capture fine-grained details of the workings of the system under study, and their ability to display emergent properties of the system—make them difficult to use for policy analysis.

ABMs describe a recursive system. At the end of each period, the system has a current state and a current behavior for each individual. Those determine, perhaps probabilistically, the next period's state. The new state and each individual's behavior determine a new behavior, and so on. The annex contains a formal description of the system. The model is described by an initial (distribution of) state(s) and an initial (distribution of) behavior(s) for each individual. Statisticians observe a run of the system, or perhaps some particular statistics, functions of the history of states and behaviors. We refer to what they see as an "observable." The model generates a distribution of observables.[13]

The value of a structural model is in its extrapolative abilities. Suppose that it is known with high confidence, after observing data from one

[12] For instance, when the capital/output ratio in an economy increases, the return to capital decreases. The share of capital in national output increases or decreases depending on the product of the two terms. Whether that share increases or decreases depends on whether a parameter of production processes, the *elasticity of substitution*, is above or below 1. This parameter measures how easily capital can be substituted for labor in production. Microeconomic studies typically find the number to be less than 0.75, but matching the macroeconomic data requires it to exceed 1. DSGE modelers choose descriptive output validation over input validation; this has implications for the predicted distribution of income.

[13] The derivation is described formally in the annex to this appendix.

environment, that p^* is the correct vector of parameter values. In any other environment for which we believe that the model still holds—Marschak's (1974) stability criterion—the parameter values p^* can be used to simulate the behavior of different policies, including policies that were not tried in the original environment if it is believed that the parameter values still apply.

The first problem that one confronts in using structural models is *identification*. Simply put, the identification question asks, Can one infer the parameter values from the observables distribution? Formally speaking, it asks whether the map from parameter values to the distribution of observables is one-to-one. That is an important question because policies will have different effects depending on the parameter values that describe an environment, and therefore one's ranking of policies is parameter dependent. At a minimum, one would like to divide the parameter space into regions that favor different policies and then determine which region best describes the world. Simple ABMs pose no unusual identification problems. For example, in the simple SIS model, up to a change in time scale, the stochastic behavior of the model is completely described by the ratio β/γ, and things that we might measure, such as the number of susceptibles or infecteds at a given time t, are stochastically increasing or decreasing, respectively, in this parameter. In complex ABMs like those created by Hoffer et al. (2009) and Eubank et al. (2004) and Toroczkai and Eubank (2005), in contrast, the parameter-identification problem is often formally unsolvable, and the most that one can learn about identification through simulation exercises is "so far, so good."

ABMs are nonlinear models; *highly nonlinear* is the usual term. Deterministic nonlinear models have three characteristics that make them difficult to use statistically:

- Sensitive dependence on initial conditions: A map exhibits sensitive dependence on initial conditions at x_0 if there is some distance $d > 0$ such that no matter how close to x_0 another initial point x'_0 is chosen, x_t and x'_t will eventually be at least distance d apart.
- Complicated limit dynamics: The limit behavior of typical linear dynamic systems is simple: Either the system converges to a steady state from any initial conditions, or they blow up, diverging to infinity. Nonlinear dynamical systems have much more complicated dynamics, including stable limit cycles of various periodicities, strange attractors, and chaotic behavior.
- Sensitive dependence on parameters: Small changes in parameter values can lead to abrupt changes in the qualitative character of system dynamics.

These properties are characteristic of deterministic systems but adding randomness does not make things simpler. The complexity of ABM dynamics can make identification difficult. In a truly complex model, one would have to simulate with the ABM across the entire parameter space to trace out the different observables distributions, and this is often not practical.

PSEUDOCOMPLEXITY

There is often less to ABMs than meets the eye. ABMs are often lauded for their ability to encode more realistic models of human behavior than do analytic social-science models, but this is a canard. Often, the agents of an ABM are described by a few possible states, and they can interact in only a few ways. For some purposes, that is a virtue. In his short story "On Exactitude in Science," Borges (1998) illustrates the problems of models as complicated as the world that they represent and the ultimate fate of such models. The moral of the story is that too much complexity is not good, and that alternatively, reductionism, an epithet often thrown at modelers by scholars of a more descriptive bent, is good. Furthermore, the goal of many scientists who develop ABMs is to demonstrate that a small number of universality classes collectively describe the behavior of (nearly) all complex systems. They see this simplicity as a virtue. Compared with conventional economic models (both analytic and computational), ABMs make individuals simpler and their interactions more complex; again, focusing on the interactions is the goal of many ABM developers.

Another way in which ABMs can be insufficiently complex is that in focusing on the details of a single system, they are too closed. For example, the heroin model of Hoffer et al. (2009) studied the workings of only one spatially contiguous heroin street market. One presumes that there are other places to buy heroin in Denver and that buyers can make choices about where to shop. That does not affect the utility of the model for studying the behavior of a single market. Movement in and out of the market is captured in a reduced-form way. But a model designed for policy analysis would have to consider the alternative venues and perhaps also the spontaneous emergence of new venues. Scaling the model up will be conceptually more accurate but will magnify all the difficulties already discussed.

PSEUDOSIMPLICITY

The point of an ABM is often to demonstrate that complicated emergent behavior is a consequence of simply described systems. Tracking emergent behaviors, however, imposes the costs of less efficient estimation and of decreased model credibility. Imagine a more complicated version of the SIS ABM, something like the models of Eubank et al. (2004) and

Toroczkai and Eubank (2005). With the right parameter values, the right structure, this model can describe the stochastic process of disease in great detail, but for policy purposes we may be interested only in the joint distribution of duration time and the total number of infections. The extra detail is unnecessary and wastes inferential power. "Knowledge," begins Marschak (1974, p. 293), "is useful if it helps to make the best decisions." ABMs have the power to deliver much useless information. Furthermore, mismatching the data in some dimensions makes the forecasts less credible in our eyes even when those dimensions are irrelevant to the policy-analysis exercise at hand.

ESTIMATION

There seem to be two general approaches to estimating parameters of ABMs. Following conventional method-of-moments techniques, one can search across the parameter space with the goal of minimizing the distance (measured according to a prespecified metric) between moments or other statistics of the simulated observables distribution and those of the empirical distribution from the data. More recently, various statistical learning and data-mining techniques have been suggested. No technique, however, has a guaranteed recipe for fitting a model on multiple time and spatial scales. The complex behavior of ABMs that makes them so useful for uncovering hidden possibilities in system dynamics stands in the way of their use for the kind of structural estimation necessary for comparing the performance of alternative policy choices.

Suppose, however, that one is confident that the data are sufficiently rich and the model sufficiently expressive to get a good fit on the multiple scales of interest in an ABM, say, by some minimum-distance estimation technique. Suppose, too, that parameters are identified and that the estimated parameter values are consistent with one's knowledge of the micro-behavioral processes. Would one then have confidence in the estimates? A basic criterion for estimator quality is the property of consistency: As the dataset grows large, estimates converge in probability to the parameter values that actually describe the data-generating process. Two consistency questions arise in models with social networks. First, consider the study of a phenomenon among unconnected communities, each with its own social network. Although the behaviors of individuals within a given community are not independent of one another, behaviors of individuals in different communities are. In such models, consistency has to do with the behavior of estimates as the number of communities becomes large; the unit of analysis is the community, and independence among communities allows for the application of standard laws of large numbers. Second, consider the study that involves ever-larger samples from a single social network. One must look to laws of large numbers for dependent random variables to address consistency, and the answers will depend on how the network grows. There are simple examples of social-interaction models in which, when the network gets large in particular ways, behavioral parameters cannot be consistently estimated.

The statistics literature does not have many consistency results for models in which the data-generating process varies discontinuously with the parameter values, but discontinuities are to be expected with ABMs for two reasons. First, the aforementioned sensitivities of long-run behavior to initial conditions and parameter shifts appear as discontinuities of, for example, stationary distributions with respect to parameters. Second, many ABMs designed to model social systems have discontinuities built in. Agents, for example, are modeled with threshold effects. Even in models in which discontinuity appears only in the large-numbers limit, as is commonly the case in models that make use of random graphs, the behavior of the model with respect to parameter changes for large but finite N can be so fragile as to make inference from parameter estimates suspect. One can often prove that parameter estimates are consistent in models that have discontinuities, but such proofs rely on other properties, such as monotonicity, that may not be present. There are few general principles to apply, and any asymptotic analysis of a given ABM may depend on fine details of the model's structure that are not readily accessible in the computational algorithm.

OTHER PATHS

My theme, in summary, is that the very complexity that makes ABMs useful for exploratory analysis creates difficulties when the task is to pin down the nature of the actual environment sufficiently to determine good policies. Are there better ways to go? Under what circumstances would a simpler model perform better? It would be hard to lay down criteria for model choices. Data availability is not a reliable guide. Some say that ABMs serve well when there is a paucity of data, but every parameter value that must be assumed rather than estimated reduces the plausibility of the model's output. Others say that ABMs serve well when there is a great deal of data, but models that are capable of generating complex output patterns may focus attention on irrelevant patterns in the data to the detriment of patterns important for the policy decision.

Although there is no rule for choosing the best model, the adage by Marschak (1974) about useful knowledge is a guide to model construction.[14] Rather than starting with a description of the data, the modeler

[14] Putnam (1974) turns this idea into a normative principle of model construction. Good models are not necessarily accurate or correct or survivors of falsification attempts; they are useful.

should start with what needs to be known and work backward to the data needs or to how the available data can be optimally deployed.

Decision theory is a guide to model construction. The decision-theory paradigm requires the planner to identify a set of feasible policies and a welfare function for evaluating them. The planner knows or forecasts the welfare that will be achieved by every feasible policy and thus can choose the best. When knowledge is incomplete, information in the planner's possession will guide them in forecasting returns, so the planner's choice will be determined by the observed information. Finally, the decision problem itself can help in identifying any needed estimation procedures inasmuch as inefficient use of data will lead to suboptimal policy recommendations. Savage's (1954) famous *Foundations of Statistics* was, in the end, the application of decision theory to determine optimal statistical procedures; the development of expected utility was but a means to an end.

Decision theory is the economist's lever. For example, the traditional economic model is expected welfare maximization, which makes use of probabilistic forecasts generated by the structural model and accounts for parameter uncertainty by positing an a priori subjective probability distribution on the set of parameters, which can be revised in light of new information. Bayesian expected utility principles are not the only way to make use of a decision-theoretic model.[15] Manski (2010) is a good place to see how decision theory is used in practice. He discusses a decision-theoretic approach to optimal policy choice without committing to subjective probability judgments about parameter values.[16]

How one deploys decision-theoretic techniques for policy analysis is a topic for another paper. The point I wish to conclude on is that a good modeling strategy strives for minimality—as simple a model as one can get away with. ABMs celebrate complexity. They are good at demonstrating what is possible but bad at pointing out what is probable. Decision theory is a guide to useful model construction because it provides a set of tools to determine how much simplicity researchers can get away with. Optimally tuning simplicity is the key to building good models for policy analysis.

[15] Milnor (1954) discusses an entire zoo of decision rules and characterizes them in terms of their different properties. A modern take on this is Stoye (2012).

[16] See also Manski (2011).

Appendix B Annex

FORMAL DESCRIPTIONS OF COMPLEX ADAPTIVE SYSTEMS

Formally, a CAS is just a mathematical function that assigns to inputs probability distributions on outcomes. The function also depends on parameters; by varying the parameters, the modeler can change some of the details of how the CAS works. We are used to thinking of parameters as scalars that indicate rates, bounds, means, variance, and so forth, but they need not be. For example, different behavioral rules that could be assigned to agents in a CAS or different rules for agent interaction are parameters of the model. The behavior of a CAS's observables can be thought of as a function whose domain is a set X of deterministic inputs, whose output is a probability distribution on a set O of outcomes, and that can be "tuned" by varying some parameters whose possible values are contained in a set P. A CAS can be described in mathematical notation this way:

$$F : X \times P \to \mathcal{O}. \tag{3}$$

That is, F is a function that maps a set X of deterministic inputs and a set P of parameters into the set $\mathcal{O}$ of probability distributions on a set O of outputs. Note that the unobservables of the system, such as the states of individual agents, do not appear directly in this description. They generate the randomness that is captured in the probability distributions that F produces; if everything is observable, the probability distributions will be trivial—the system is deterministic.

Of course, not all possible functions qualify as CASs. CASs (and the ABMs that implement them) are more specific than that. At the greatest level of generality, they are examples of *random dynamical systems with complete connections*.[17] The randomness comes from hidden state variables, labeled S in Figure B-2. Figure B-2 describes the general structure of a typical CAS. Wavy lines represent stochastic dependence, and solid lines represent deterministic functional dependence. Thus, o_{t+1} is a function of o_t, and s_{t+1} is a random variable whose distribution is determined by o_t and s_t. The symbol π stands for the conditional probability of s_t given o_{t-1} and s_{t-1}, and f is a function that maps each o_{t-1} and s_t into o_t. The parameters p describe π and f. Even that is too general. In the ABMs that I am familiar with, the diagonal wavy lines are gone; each s_t depends only on s_{t-1} and not on o_{t-1}. Such structures are known as *hidden Markov*

[17] This mathematical construct has been important for the general analysis of adaptive learning models. See Iosifescu and Grigorescu (1990) and Norman (1972).

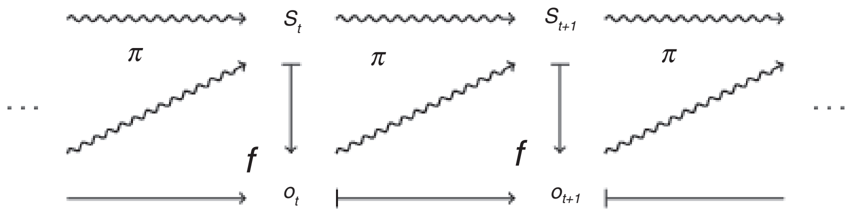

FIGURE B-2 A random dynamical system with complete connections.

models. In many cases, the hidden states are just independent random variables. It should be apparent—and this is the point of the figure and the preceding jargon—that at a formal level, an ABM is just another statistical model. Whatever is good or bad about ABMs lies in the details of their specification and implementation. Just to touch base with the high-level description, the inputs are o_0 and s_0 (which will itself be drawn from a given probability distribution), and the outcome is a probability distribution on the set of possible sequences $(o_1, o_2, \ldots, o_T)$, which is generated as described above. The parameters determine the probability distribution on outcome sequences through their effects on π and f. The o_t are high-dimensional objects that describe the current action or state of each agent. Emergent properties have to do with aggregates, such as the population mean and variance of behavior, and other less obvious functions of the observables. For example, in the SIS CAS, the extinction time is an emergent property.

REFERENCES

Borges, J. L. 1998. On exactitude in science. In *Collected fictions*. A. Hurley (Trans.). New York: Penguin Books. P. 325.

Cartwright, N. 1979. Causal laws and effective strategies. *NOÛS* 13:419–437.

———. 1989. *Nature's capacities and their measurement*. Oxford, UK: Oxford University Press.

Eubank, S., H. Guclu, V. S. A. Kumar, M. V. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang. 2004. Modelling disease outbreaks in realistic urban social networks. *Nature* 429(6988):180–184.

Halpern, J. Y., and J. Pearl. 2005a. Causes and explanations: A structural-model approach. Part I: Causes. *The British Journal for the Philosophy of Science* 56(4):843–887.

———. 2005b. Causes and explanations: A structural-model approach. Part II: Explanations. *The British Journal for the Philosophy of Science* 56(4):889–911.

Heckman, J. J. 2000. Causal parameters and policy analysis in economics: A twentieth century retrospective. *The Quarterly Journal of Economics* 115(1):45–97.

Hoffer, L. D., G. Bobashev, and R. J. Morris. 2009. Researching a local heroin market as a complex adaptive system. *American Journal of Community Psychology* 44(3-4):273–286.

Holland, J. 1975. *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control and artificial intelligence*. 2nd ed. (1992). Cambridge, MA: MIT Press.

Hood, W. C. and T. C. Koopmans (Eds.). 1953. Studies in Econometric Method. Cowles Commission Monograph 14. New York: Wiley.

Hume, D. 1777. *Enquiries concerning the human understanding, and concerning the principles of morals*. Edited by L. A. Selby-Bigge. 2nd ed. (1902), Oxford, UK: Clarendon Press.

Iosifescu, M., and S. Grigorescu. 1990. *Dependence with complete connections and its applications*. Cambridge, UK: Cambridge University Press.

Lewis, D. 1973. Causation. *The Journal of Philosophy* 70(17):556–567.

Lucas, R. 1976. Econometric policy evaluation: A critique. In *The Phillips Curve and labor markets*, edited by K. Brunner and A. H. Meltzer. Amsterdam, Netherlands: North-Holland Publishing Company. Pp. 19–46.

Manski, C. F. 2010. Vaccination with partial knowledge of external effectiveness. *Proceedings of the National Academy of Sciences of the United States of America* 107(9):3953–3960.

———. 2011. Choosing treatment policies under ambiguity. *Annual Review of Economics* 3:25–49.

Marschak, J. 1974. Economic measurements for policy and prediction. In *Economic information, decision, and prediction,* Vol. 7–3. Netherlands: Springer. Pp. 293–322.

Milnor, J. 1954. Games against nature. In *Decision processes*, edited by R. M. Thrall, C. H. Coombs, and R. L. Davis. New York: Wiley. Pp. 49–60.

Norman, M. F. 1972. *Markov processes and learning models*. New York: Academic Press. http://public.eblib.com/choice/publicfullrecord.aspx?p=453098 (accessed November 17, 2014).

Pearl, J. 2009. *Causality: Models, reasoning, and inference*. 2nd ed. Cambridge, UK: Cambridge University Press.

Putnam, H. 1974. The "corroboration" of theories. In *The philosophy of Karl Popper*, edited by P. A. Schilpp. La Salle, IL: Open Court. Pp. 221–240.

Savage, L. 1954. *The foundations of statistics*. New York: Wiley.

Schelling, T. C. 1971. Dynamic models of segregation. *Journal of Mathematical Sociology* 1(2):143–186.

Simon, H. A. 1953. Causal ordering and identifiability. In *Studies in econometric method*, edited by W. C. Hood and T. C. Koopmans. New York: Wiley. Pp. 49–74.

Stoye, J. 2012. New perspectives on statistical decisions under ambiguity. *Annual Review of Economics* 4(1):257–282.

Tesfatsion, L. 2015. *Verification and empirical validation of agent-based computational models*. http://www2.econ.iastate.edu/tesfatsi/empvalid.htm (accessed November 17, 2014).

Tinbergen, J. 1968. *Statistical testing of business-cycle theories*. New York: Agathon Press.

Toroczkai, Z., and S. Eubank. 2005. Agent-based modeling as a decision-making tool: How to halt a smallpox epidemic. *The Bridge* 35(4):21–27.

Wright, S. 1921. Correlation and causation. *Journal of Agricultural Research* 20(7):557–585.

———. 1925. Corn and hog correlations. *U.S. Department of Agriculture Department Bulletin* 1300:1–60.

Young, H. P. 2001. *Individual strategy and social structure: An evolutionary theory of institutions*. Princeton, NJ: Princeton University Press.

Appendix C

Assessing Agent-Based Models for Regulatory Applications: Lessons from Energy Analysis

Alan H. Sanstad[1]

ABSTRACT

Agent-based modeling (ABM) has been proposed as a promising method for analyzing behavior patterns related to smoking in connection with regulation of tobacco products. This possibility raises questions about model validation and evaluation, uncertainty quantification, and how the models should be judged with respect to suitability for policy applications. These issues have long been present in social science–based and policy-focused computational modeling; computational energy modeling is an important example. This paper reviews energy modeling from methodological and epistemological perspectives to draw lessons for ABM regarding model validity, the treatment of uncertainty, and criteria for decision makers to apply when considering agent-based models for use in regulation.

INTRODUCTION

Computational modeling of social and economic systems and behavior developed in the 1950s and 1960s and became well established in the 1970s. In this category of modeling, so-called agent-based modeling (ABM) has emerged as an active research field.[2] Recently, ABM has been proposed as a promising method for analyzing smoking-related behavior patterns, for use in regulating tobacco products.

---

[1] Berkeley, California. alansanstad@gmail.com.

[2] Throughout this paper, *ABM* will be used as an abbreviation of both "agent-based modeling" and "agent-based model(s)."

Such applications of ABM raise basic questions that have a long history in social science–based and policy-focused computational modeling, including

- What can be learned from computational models about behavior or other phenomena of interest?
- How can the uncertainty associated with their structure and inputs and their quantitative outputs be analyzed?
- How should their validity and or utility—that is, their usefulness—be evaluated in a regulatory context?

Regarding ABM specifically, here as in other contexts, some claims have been made about its capabilities and putatively unique advantages over other modeling approaches, including its ability to represent social and behavioral phenomena with a high degree of detail, and the model fidelity that this is said to provide. How should these claims be assessed? More generally, does ABM have characteristics that pose validation or evaluation questions that are different from ones that are relevant for other types of computational social science models? Or does the nature of ABM somehow obviate such considerations?

Such questions are fundamentally epistemological: They pertain to defining and characterizing what knowledge can be generated by computational models and to understanding how their outputs should be applied to decision making. Notwithstanding both their importance and the attention devoted to them in different social science and policy modeling domains over many years, those questions continue to be challenging and generally unresolved.

The issues are well illustrated in the field of computational energy modeling. Since the 1970s, this type of modeling has steadily expanded in scope and in importance for energy regulation and policy making. In recent years, in addition to energy analysis specifically, it has increasingly been applied to the problem of mitigating emissions of carbon dioxide (CO_2) from energy production and consumption. In terms of the level and range of activity, prevalence, and influence, it is perhaps the primary example of computational policy modeling.[3] Epistemological issues were recognized early in energy modeling's more-than-four-decade history and continue to be important. The premise of the present paper is that, notwithstanding technical and methodological differences, energy modeling can provide valuable insights and lessons for ABM with respect to model validity and evaluation, uncertainty quantification, and policy applications.

[3] Energy modeling is also an academic field; however, the preponderance of such work deals at least implicitly with policy and regulatory applications rather than constituting "basic" research.

The paper is organized as follows. Immediately below, key terminology is explained. A brief overview of energy modeling is presented next. The paper then turns to a discussion of methodological and epistemological issues, particularly model calibration and its relation to uncertainty, and the value of increasing model complexity. Examples of agent-based modeling of energy systems are then presented, and two particular studies of this type discussed in detail. The concepts of fundamental model uncertainty and robustness and their potential relevance to ABM are then briefly reviewed. Lessons from energy modeling that are applicable to ABM are followed by recommendations for assessing potential regulatory applications of ABM and concluding remarks.

Terminology

In the context of computational modeling, such terms as *validity* and *validation* have not only different technical meanings but different connotations among and in some cases within disciplines. Particularly in the social sciences, the terms may be interpreted as representing concepts and methods that are more appropriate to the physical sciences. In energy modeling specifically, validity and validation as such are not only not generally discussed or practiced, respectively, but in some quarters are viewed as fundamentally inapplicable. There are, however, no standard or generally accepted alternatives either conceptually or in nomenclature. Thus, in this paper, the terms *validity* and *validation* will be used, as well as *evaluation*, and *quality.* But the reader should understand that these terms are highly approximate and simply provide a shorthand for discussing the assessment of computational models and their usefulness in applications.

Energy models here refers to computational models, based on economic and optimization principles, of energy systems, entire economies with particular detail on energy sectors, or specific energy-using sectors, particularly residential and commercial. Technically, the system and economic models are generally (although not exclusively) of the equilibrium type (represented by systems of nonlinear algebraic equations), the mathematical programming type, or the optimal control type.[4] "Partial equilibrium" models in this case represent energy demand and supply sectors and their market interactions. "Computable general equilibrium" (CGE) models also have similar components of energy demand and supply sectors and their market interactions, but these components are embedded in a full representation of a complete economy, which among other features includes direct or indirect interactions between the energy sector and all other parts of the

[4] In economics, these types are not mutually exclusive, but the distinctions among them are useful in understanding the contemporary landscape of energy modeling.

overall economy. Many models of the electric power system are of the linear programming type.

In this context, *computation* means finding solutions to the optimization problem or the system of equations represented by the model.

An important development in the last several decades has been the advent of "integrated assessment" models for the economic and policy analysis of global climate change. Those are energy models coupled to reduced-form representations of the climate system, and in some cases, other physical and ecologic systems.[5] In what follows, for brevity's sake the term *energy model* will generally be used rather than *energy and/or integrated assessment model*; this will not distort the discussion or the conclusions, but the reader should be aware of the distinction.

THE CURRENT LANDSCAPE OF ENERGY MODELING

This section discusses and gives examples of non-agent-based energy models and their policy applications.

Energy models are not just widely applied to but have become the predominant analytical methodology for energy policy and regulatory analysis in the United States. Although in this, as in other computational modeling applications, the models are often referred to as "tools," such a characterization understates their role and influence. In fact, to a great degree they define the universe of policy discourse and determine what questions can be asked, what form answers take, and what constitute useful data by virtue of providing model inputs.

Energy models are used by agencies and other policy and regulatory bodies at all levels of government. At the federal level, the Energy Information Administration (EIA) maintains and applies the National Energy Modeling System (NEMS), which projects the evolution of the U.S. energy system over several decades (EIA, 2009; NRC, 1992).[6] The primary use of NEMS is in the production of the *Annual Energy Outlook* (*AEO*), which documents the details of each year's updated projection, including details on individual fuels (electricity, natural gas, and so on), energy production, and energy consumption in different sectors. (An example is presented below.) It is also used to analyze the potential effects of proposed policies

[5]The best-known example is Nordhaus's DICE (Dynamic Integrated Climate Economy) model, which is based on optimal control principles (Nordhaus, 2008).

[6]EIA is the federal agency that has principal responsibility for analyzing energy issues. Its activities include the collection and dissemination of energy statistics and computational energy-policy modeling, both short-term and long-term and on national, international, and regional scales. Although in the U.S. Department of Energy, EIA reports to Congress, for which it conducts analyses of energy topics on request.

on the energy system, including effects on supply, demand, prices, and costs.[7]

In applications of this type, NEMS and other models are used in a standard analytical structure: the model is solved for a "reference" or "baseline" or "business-as-usual" projection without the policy in question, and then re-solved with the proposed policy introduced into the model (for example, by introducing an emissions tax or a policy-induced technological improvement). The results are then compared to identify the policy's effects as represented by the model. Both types of projection are called *scenarios*. An example is shown in Figure C-1, which displays the output of a NEMS–*AEO* reference case, specifically, how much electricity is generated from different sources, out to the year 2040. (The vertical axis unit is trillions of kilowatt hours.) The "History" period is based on empirical observations and "Projections" on model output.

A recent example of the use of NEMS for policy analysis is a study of the proposed Clean Energy Standard Act of 2012, which aimed to reduce greenhouse gas (GHG) emissions from electric power generation (EIA, 2012).[8]

The U.S. Environmental Protection Agency (EPA) uses several types of policy models for energy regulatory analysis. It uses two CGE models for analyzing environmental policy problems related to energy: the Applied Dynamic Analysis of the Global Economy (ADAGE) model (Ross, 2005), and the Intertemporal General Equilibrium (IGEM) model (Goettle et al., 2007). Whereas NEMS was developed and is maintained and operated by EIA itself, these two are proprietary models developed outside EPA but used by the agency under long-term contracts. An example of their use was an analysis of the proposed legislation HR 2454 of 2009, colloquially known as the Waxman–Markey Act, which called for the establishment of a GHG emissions "cap and trade" regime and other emissions-reducing measures to reduce U.S. emissions to more than 80 percent below their 2005 level by 2050 (EPA, 2009).[9] EPA also uses a commercial proprietary linear programming model of the electricity system (EPA, 2010) to analyze pollution emissions; a current application is the development of regulations to reduce CO_2 emissions from power generation.

At the state level, some energy regulatory entities use modeling for such purposes as utility regulation, energy-efficiency policy analysis, and,

[7]NEMS is a partial-equilibrium model and represents the interconnected system of markets that links energy supplies and demands driven by economic, demographic, and other factors over about a 3-decade horizon with annual steps.

[8]The study was requested by Senator Jeff Bingaman (D-NM), chairman of the Senate Committee on Energy and Natural Resources at the time.

[9]The study was requested in 2009 by Henry Waxman (D-CA), chairman of House Energy and Commerce Committee, and Edward Markey (D-MA), chairman of the Energy and Environment Subcommittee at the time.

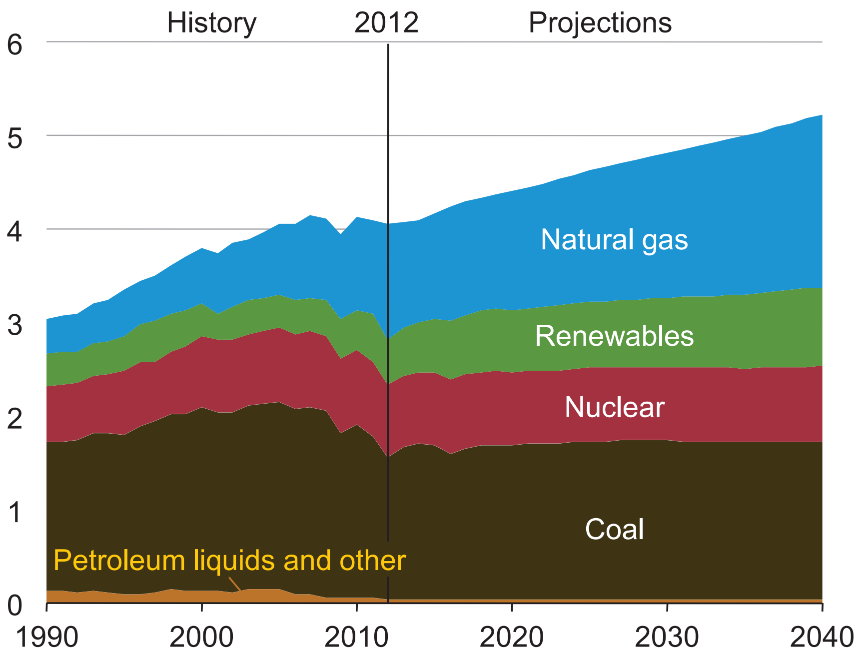

FIGURE C-1 NEMS projection of energy use for electric power generation, in trillions of kilowatt hours by year.
SOURCE: EIA, 2014.

in recent years, GHG emissions abatement (e.g., Goldstein et al., 2008). Modeling also plays a crucial role in operations management of and planning for the electric power transmission system. Outside the government, political advocacy groups have increasingly turned to modeling to develop and lobby for (or against) specific energy-related and environment-related proposals for legislation. For example, using a version of the linear programming model mentioned above, the Natural Resources Defense Council developed a policy architecture for electricity CO_2 emissions abatement (Lashof et al., 2013).

Methodological Aspects

Although, as noted above, energy models are for the most part built on a small number of underlying principles, there is a great deal of variety among models across the entire field; to some extent, this reflects the many ways that these principles can be implemented in practice. However, a small number of models can be considered predominant and most influential in the United States. Broadly speaking, they share several critical features. First, they are deterministic: There is no explicit representation of uncertainty in the model on the part of either the model builder or the agents—consumers and firms.[10] Second, they are high-dimensional; that is, they contain a great deal of detail regarding elements that, depending on the model, may include production, consumption, technologies, economic sectors, markets, policies, and so forth.

With few exceptions, energy models designed for and applied to long-run—multidecade—analysis are parameterized by calibration.[11] The meaning of "calibration" varies among fields of computational modeling. In energy modeling, the essential aspect is that values of model parameters are set *without the direct use of statistical methods or techniques relating the model to empirical data.* Some models contain simplified engineering descriptions of individual technologies or technology types, and parameter values are obtained from, for example, engineering studies or surveys. In those and in models that have a more aggregate structure, there are two primary calibration techniques. First, most model parameters are taken from other sources. A key example is "substitution elasticities," which characterize how consumers and firms make trade-offs in their choices among goods and services. These elasticities have often been adopted from empirical studies that did use statistical methods but typically applied to econometric (statistical) models that have considerably different structure from and simpler structure than numerical energy models.[12] In the common circumstance in which a range of estimates appears in the literature, the mean estimate is often used.

Second, parameter values can be set by "tuning." The most important examples in energy modeling are parameters determining the magnitude of aggregate improvements in energy productivity, which are directly analogous to the labor-productivity parameters that are common in macroeconomic models. The magnitudes of the values of these energy-productivity parameters are commonly determined by exogenously setting them informally (i.e., without using a statistical method) to reproduce approximately the historical observed trends in the ratio of economic output to energy

[10] This is an oversimplification in that what could be called experimental uncertainty quantification has been conducted with several of the models in question. However, their basic and typically used mode is deterministic.

[11] The qualification "long-run" is used to distinguish them from modeling conducted in the electric utility industry for real-time or 1-day-ahead forecasting and optimization, which increasingly incorporates explicitly statistical and stochastic components and techniques.

[12] A rare exception is the IGEM model mentioned above, in which case many of the parameter values are obtained econometrically directly, that is, by an integrated estimation of the equilibrium model itself with suitable empirical data.

input, whether at the level of an individual sector or of the entire economy (the ratio of gross domestic product to economywide energy consumption).

CURRENT STATE OF THEORY AND PRACTICE

Energy models are subject to widely varying degrees and types of assessment. Those used by some federal agencies, for example, are subject to review requirements. However, there are no specific, generally applied theoretical, empirical, or computational methods or procedures for defining or determining energy model validity, verisimilitude, or quality. This in part reflects a sustained inattention to epistemological issues after a period in the 1970s and 1980s during which they received considerable attention.[13] However, it also has to do with the special challenges of addressing such issues for models designed for and applied to projections of an energy system or economy decades into the future. Even if such a model is in some fashion empirically grounded in current or historical data, it is not clear whether or how such grounding provides evidence or assurance of its suitability for such long-run analysis. Indeed, it is difficult even to define *suitability* or *validity* in this circumstance. Moreover, there is prima facie a level of uncertainty associated with multidecade projections that is categorically different from that involved in, for example, short-run economic forecasting—up to several calendar quarters.

In practice, broadly speaking, model quality is claimed on such grounds as internal consistency, the plausibility of assumptions and results, and the usefulness of generating "insights," whereas uncertainty is addressed by appealing to the logic of scenarios. The following subsections discuss those topics in more detail.

Calibration and Knightian Uncertainty

Returning to the energy-modeling overview and scenario example (from the NEMS model) of the previous section: How should such projections be interpreted? According to EIA (2014, p. iii),

> projections by EIA are not statements of what will happen but of what might happen, given the assumptions used for any particular scenario. . . . Energy models are simplified representations of [the energy system]. Projections are highly dependent on the data, methodologies, model structures, and assumptions used in their development. Behavioral characteristics are indicative of real-world tendencies rather than representations of specific outcomes. . . . [These] projections are subject to much uncertainty. Many . . . events that shape energy markets are random and cannot be anticipated. In

[13]This history is sketched in the annex to this appendix.

> addition, future developments in technologies, demographics, and resources cannot be foreseen with certainty. Many key uncertainties in the *AEO2014* . . . are addressed through alternative cases.[14]

This statement is a candid acknowledgment of fundamental modeling limitations and a straightforward and reasonable disclaimer. However, it also expresses a perspective on uncertainty and scenario—and implicitly model—validity that, although written regarding NEMS and the *AEO*, captures much of current standard energy and integrated assessment modeling epistemology more generally. This is reflected in methodological views expressed by other modeling groups. Clarke et al. (2007), for example, state, "Model-based scenario analysis is designed to provide quantitative estimates of multiple outcomes and to assure consistency among them that is difficult to achieve without a formal structure" (p. 43). They also suggest that "[a]lthough the future is uncertain and the scenarios are strongly dependent on many underlying assumptions, this research provides useful insights for those engaged in climate-related decision making" (p. 5). Similar views are expressed by other important stakeholders, such as the Intergovernmental Panel on Climate Change (2000, p. 3):

> Scenarios are alternative images of how the future might unfold and are an appropriate tool with which to analyse how driving forces may influence future emissions outcomes and to assess the associated uncertainties. . . . The possibility that any single emissions path will occur as described in scenarios is highly uncertain.

Although not so intended, statements of this type are in a way implicit explanations of the general absence of formal, quantitative uncertainty analysis and validation in energy modeling. An important reason is the calibrationist methodology within which the models are developed and applied. Dawkins et al. (2001) discuss calibration in considerable detail; although their focus is macroeconomic modeling and applied general equilibrium modeling especially in international trade applications, their observations apply to energy modeling as well (including both computable general equilibrium and other types):

> [M]odellers typically see their simulations largely as numerical implementations of theoretical structures. To them, the widespread use of a particular structure in the theoretical literature is an indication of its worth, so that they seek less to test or validate models and more to explore the numerical implications of a particular model, conditional on having chosen it. . . . The focus of micro modellers is to generate insights about the effects of

[14]The alternative cases are a small number of scenarios that assume different economic growth rates, world oil prices, or technologic progress.

> policy or other changes conditional on a particular theoretical structure, rather than to test theory itself. (Dawkins et al., 2001, p. 3672)[15]

That models are deterministic and their outputs conditional does not, of course, imply that they do not contain and reflect uncertainty. However, in the calibrationist approach to energy modeling and the models' use to generate scenarios, such uncertainty is of a type often referred to as Knightian; it cannot be described or quantified by assigning probabilities.[16] In energy modeling, in principle, the scenario approach is a reasonable way of addressing Knightian uncertainty. But the common practice of computing small numbers of reference cases or scenarios and policy scenarios that are incremental to the reference cases in effect suppresses a great deal of uncertainty.

The reason is that it is rarely the case that the particular set of input parameter values chosen for the underlying model is *uniquely* justified.[17] The evidence often suggests no more than that the appropriate magnitudes of specific parameters, such as substitution elasticities governing economic choices among different commodities, probably fall within a particular range. However, it is customary for modelers to choose, for example, the midpoints of the ranges for the values of the parameters. In deterministic models, this is not justified as, say, the mean of a uniform distribution; rather, it is simply a heuristic. But in such a case, the available information implies that every point in the *entire* interval is equally plausible for use in the model. For a given model, there may be multiple such "equiplausible" intervals for various parameters. Under those circumstances, the complete warranted input set constitutes all possible selections of specific parameter values from this set of intervals. (For example, if there are two uncertain parameters and two such intervals, the entire input set is a rectangle.) Thus, all the simulations that would in principle be generated by running the model for all of these selections of parameter values are equally "valid" or plausible. Even when particular scenarios are defined by, for example, specific assumptions regarding policy—such as the magnitude of a CO_2 emission tax—this implies that taking account of all the relevant information requires that the potential effects of the policy should be computed by running the model across the entire input set defined above.

That, with a small number of exceptions, this is not done in standard energy modeling practice is what was referred to above as a suppression

[15] These observations are about "applied general equilibrium" economic modeling, which is the application of microeconomic principles to the representation of an entire economy. Hence the use of the term *micro*-.

[16] This distinction was introduced by the economist Frank Knight (1921).

[17] This is pointed out, for example, by Clarke et al. (2007) in their discussion of reference-case scenarios in global integrated assessment modeling.

of uncertainty, in this case Knightian. In effect, the conventional method of choosing a single set of representative inputs means that scenarios are, abstractly, single points in a very high-dimensional space of equally credible points—that is, the hypothetical space of all model simulations, indexed by parameter value choices.[18]

In calibration-based modeling, parameter sensitivity analysis is sometimes applied to address model or scenario validity putatively, whether implicitly or explicitly. Specifically, a finding of low sensitivity of model output to small changes in the value of one parameter or a small number of parameters is claimed as evidence that the base output is in some sense valid. With calibration, however, that logic is flawed. Conceptually, sensitivity analysis is "intrinsic" in the sense that, essentially by definition, it provides information about the model itself, not about the relationship between the model and the system or phenomenon that it represents. When a model can actually be demonstrated to be invalid—to be in some way an inaccurate representation—a finding of low sensitivity merely indicates that it is, in a manner of speaking, robustly inaccurate around the base parameter values.

However, a similar point holds in contexts like the one discussed here, in which validity itself is difficult to define or measure but a model has recognized credibility, usefulness, or acceptable quality. A given calibrated model with chosen parameter values may provide useful information to a decision maker. In this case, a finding of low sensitivity to the values is itself useful insofar as it suggests that the model may be robustly informative around the values. Nevertheless, that cannot be claimed to demonstrate any form of validity in the sense of empirical fidelity, nor would a finding of high sensitivity necessarily demonstrate invalidity.

Even in calibrated models, there may be cases in which uncertainty regarding a parameter value is either absent, on the one hand, or can be represented by a probability distribution (that is, is non-Knightian), on the other. In some instances, the available evidence may support the specific parameter values that are chosen as opposed to suggesting only a range of equally plausible values. If the exact values themselves are known with certainty, sensitivity analysis provides no information, because nearby values cannot occur. It is in the more common circumstance, in which the values are known to have a most likely but not certain value, that sensitivity analysis may be most informative. In such cases, however, the implication is that some sort of probability distribution can be assigned to the parameter (as opposed to its value being characterized by Knightian

[18] A noteworthy exception is the work of Lempert and colleagues (e.g., Lempert, Popper, and Bankes, 2003; Lempert, Bryant, and Bankes, 2008), which will be discussed later in the paper.

uncertainty). Thus, although still informative, sensitivity analysis is an element of uncertainty quantification.[19]

Complexity and Validity

Although, as noted above, there is an absence of concepts and methods for assessing energy models objectively, contemporary energy-modeling discourse and practice reveal a widely held belief that increasing levels of model detail and complexity yield greater validity or verisimilitude and improved usefulness for policy applications. Calibration is both a cause and a consequence of the trend toward ever more complex models. On the one hand, once a calibration rather than estimation philosophy is adopted, models are to a great degree freed from constraints of the data requirements that arise when statistical techniques are used. On the other hand, that freedom encourages the addition of detail, in that model components need not be directly tied to data for purposes of parameterization. In turn, the possibility of validation in any traditional sense becomes increasingly remote.[20]

At the same time, however, without actual metrics for gauging the various dimensions of model quality, there are no formal grounds for asserting that increasing detail and complexity yield greater validity. But greater detail inarguably results in a larger and more complex system of relationships between model inputs and outputs. The EIA phrase quoted above is useful in understanding the implications: "[Model] projections are not statements of what *will* happen, but [rather] of what *might* happen" (emphasis added). Putting more detail in a model increases the number of possibilities of what might happen, the complexity of any single "instance," or both. In this respect, increased detail *increases* implicit uncertainty in calibrated models.

Beyond energy applications specifically, the belief that greater detail itself necessarily increases the validity of a calibrated computational model has little or no formal theoretical or empirical justification. There is an instructive contrast with how this and related issues are addressed in statistics and information theory. First, in the latter the degree of validity—accuracy or inaccuracy—of a model can be precisely defined in terms of statistical bias, and the extent to which additional model detail or complexity re-

---

[19] This assumes that the parameter itself has previously been statistically estimated, not the model containing the parameter, so that the model itself is still of the calibration type.

[20] The steady increase in model complexity also reflects the economic consequences of radical improvements in computational hardware and software over the last several decades. Computation itself has become extremely inexpensive while software advances have made model construction much easier. At the same time, observation and measurement—empirical analysis, including that which supports model construction—have remained expensive. Thus, in effect, computation has been steadily substituted for measurement.

duces bias—increases accuracy—can be quantified. Second, the uncertainty associated with a model's representation of the underlying phenomenon, system, or relationship that it describes can be defined and quantified in terms of variance. Moreover, the relationship between these two quantities of accuracy and uncertainty can be exactly characterized; the result is known as the bias–variance trade-off: An increase in model dimensionality that reduces bias is directly associated with an increase in variance (Burnham and Anderson, 2002). Those concepts have no well-established analogues in calibration-based computational policy modeling in general (including energy modeling).[21]

EXAMPLES OF AGENT-BASED ENERGY MODELING

Energy has not been a major focus of ABM, but there is a sufficient body of work on which to base discussion of several of the key issues discussed so far in this paper.

Electricity-Market Modeling

Most ABM in energy has focused on the electric power system, in particular wholesale electricity markets. An excellent critical review of this work is provided by Weidlich and Veit (2008). As they describe, these markets have characteristics that make them challenging to analyze with conventional economic optimization and equilibrium modeling. They deviate from the conditions of the "perfect competition" assumption that underlies most such modeling, and the participants in the markets engage in "strategic behavior"; that is, in making decisions they take into account the potential decisions, and reactions, of other participants. ABM methods have been applied to study, for example, the market consequences of variation in agents' behavioral rules. Another ABM application is comprehensive modeling of overall electric power systems that incorporates high levels of detail about technology and other elements (Barton et al., 2000; Koritarov, 2004). More recently, models for large-scale energy and GHG policy analysis have been created that contain the same general components as the energy models discussed in this paper but are based on ABM principles; Gerst et al. (2013), for example, represents the global economy and energy system and is used to analyze such topics as international negotiations on emissions-abatement agreements.

Weidlich and Veit reported sparse attention to validation and verifica-

[21] Although framed in different terms, an insightful perspective on several of the issues discussed in this section was provided by the Congressional Research Service in a report on energy modeling; see annex.

tion issues in agent-based energy modeling as of the time of their review. Regarding the electricity system models specifically, they observe that "the scientific usefulness and academic contribution of large-scale [agent-based] models that integrate an enormous amount of details has not yet been proven" (2008, pp. 1750–1753). It is in addition worth noting that, to the present day, these large-scale models have not gained acceptance by regulators or policy makers.

Models of Individual Behavior

In addition to those market and system models, Weidlich and Veit (2008) note an example of agent-based energy modeling studies focused on consumer or household behavior. Ehlen et al. (2007) studied "dynamic" or "real-time" electricity pricing, which refers to retail (consumer) prices, also called tariffs, that vary over the course of one day (24 hours) to reflect the marginal cost of power generation at different times. Most residential electricity customers in the United States have tariffs (stipulating prices as a function of the level of consumption and its timing) that are uniform or "flat," not reflecting the time of use (although in some parts of the country the tariffs incorporate a "block rate" price that varies with the level of consumption). However, the marginal generation cost is typically higher during daylight hours, when, for example, air-conditioning demand is highest (in some regions and climates). The purpose of dynamic pricing is to align marginal prices with marginal costs to increase the economic efficiency of electricity markets and the operation of the electric power system.

Economists have long advocated dynamic electricity pricing, and it is the subject of a theoretical and empirical literature on the behavior of consumers who face dynamic electricity prices and on how consumer preferences regarding energy consumption interact with their technological and economic environment to determine their patterns of electricity use. However, it has been difficult to implement in practice, largely because of consumer reluctance to adopt it voluntarily (that is, to use electricity under a dynamic tariff) or to accept its imposition. In recent years, it has gained increasing attention as a means of supporting the development of the "smart grid," an electricity system that uses advanced information technology to improve operations and that incorporates advanced energy technology. For this among other reasons, understanding consumer adoption or non-adoption of dynamic electricity tariffs is an important and policy-relevant goal. The following paragraphs discuss in turn Ehlen et al. (2007) and a more recent study of the same topic.

The Effects of Adoption of Dynamic Pricing in an Electricity Market

Ehlen and colleagues use ABM to study the effects of real-time pricing on residential electricity consumption and in turn on the operation of the electricity market. Their framework comprises an ABM of residential, commercial, and industrial electricity users and a simulation model of the power-generation sector. A "demand aggregator" buys power from the generator and sells it to users, and the analysis focuses on the effects of households' choice of pricing contract on the functioning of the electricity market.

Households' hourly demand during the course of a day is divided into three categories: "optional," which is of "relatively low consequence" and can be interrupted without households' needing to "make it up," such as power for lighting and television; "moveable" use, which is of "medium consequence" and can be shifted to other times, such as dish-washing or clothes-washing; and "immoveable" use, which is highly time-sensitive, such as food refrigeration or air conditioning during hot weather. Households have a desired mix of use represented by the percentage of their preferred use assigned to each category, which is fixed (numerically) a priori. Households face either uniform or real-time pricing contracts and have a fixed budget for electricity.

Households allocate their use throughout the day with the goal of consuming their desired mix among the categories. Under uniform pricing, that is straightforward. In contrast, at the beginning of a day, a household under dynamic pricing calculates the cost of its desired use and timing for that day under the hourly prices. If the cost exceeds the budget constraint, the household reallocates its use according to a "greedy scheduling algorithm," a search-based optimization procedure. First, the movable use that is of highest cost is shifted to other, lower-cost hours; if the budget constraint is still exceeded, the next-highest-cost movable use is shifted, and so on. If this procedure does not result in a planned pattern of use that satisfies the budget constraint, optional use is curtailed.

Those calculations are embedded in a higher-level process of households' choosing, or not, to adopt real-time pricing contracts and, once they are adopted, whether to stay with the new contract or revert to a uniform tariff. The model's representation of these phenomena is justified as follows:

> Whether a particular household adopts a real-time pricing contract is a complex process dependent on at least four factors: the relative economic advantage of the alternative contracts, the [transaction] cost of initiating the change . . . the willingness of the household to experiment with the new form of power contract, and the social exposure and acceptance of the contract. . . . Willingness [to experiment] . . . at least for a subset of households, is likely a function of the current level of adoption in the

> market place. This adoption, in turn, is a function of information . . . from the market and . . . social/cultural interactions with other households. (Ehlen et al., 2007, p. 6)

The details of how these assumptions are implemented are not clear from the paper's description. It is stated both that each household is randomly assigned a probability of experimenting with real-time pricing if it has not already adopted one, and that "social networks are modeled explicitly" and affect the decision to adopt (or to stay with the new contract if the shift is made). The greater the number of existing adopters, the more likely that a household with a uniform contract will switch; conversely, the greater the number who revert to uniform, the more likely that a dynamic-tariff household will also revert.

The model simulations explore the relationships among tariff types, tariff switching, and consumers' electricity use and the implications for the functioning of the modeled electricity market, including the profitability of demand aggregation. To briefly assess the model: First, whether the adoption decision is a "complex process" and whether or how the four factors noted above affect it are empirical questions (as is whether factors other than, or in addition to, those four are relevant). However, the authors do not cite or draw on the literature on dynamic-pricing adoption by households. Economists would readily agree that the first two factors are important. But Ehlen et al. provide no rationale for the particular structure—that is, mathematical—with which costs and the economic decision rules are incorporated into the model. That "willingness to experiment" is a factor that could be considered vacuously true, inasmuch as an unwilling household, by definition, will not adopt voluntarily. In any case, however, the claims that it is "likely a function of the current level of adoption" and that adoption is in part a "function . . . of social/cultural interactions with other households" are speculative; no evidence is provided to motivate or support them.

Second, the model's representation of electricity demand and consumer choices as conditional on the tariff type is problematic. Actual households' responses to dynamic pricing—including what end-use energy services are "optional," "moveable," and "immoveable"—are functions of occupants' preferences, equipment, and costs. The quantitative measurement of consumer responses to dynamic pricing, including the effects of such a factor measured through detailed econometric (statistical) analysis (see, e.g., Borenstein, 2013, and the references therein), is part of the extant dynamic-pricing literature that is ignored by Ehlen et al. No source or explanation is given for the values of several basic parameters related to consumers' decisions: their monthly budgets, the cost of switching, and the "contract transaction cost."

Finally, the authors assert that their framework for consumer behavior has "sufficient fidelity to characterize households" (2007, p. 4) that have different characteristics and constraints of several types. But there is no discussion or demonstration of how or why the model's structure results in fidelity. Moreover, although the simulations assume 10,000 residential households, the key calibration—that of the allocation of electricity consumption among different end-use types—is carried out with *aggregate* data on residential energy use in California, that is, statewide averages of consumption among all households. Thus, what is being analyzed is not a large number of truly heterogeneous agents, which is implicitly part of the basis of the claim of "fidelity," but rather a large number of *representative* agents.

Analyzing Adoption of Dynamic Pricing

A second example is another ABM-based study, by Kowalska-Pyzalska et al. (2014), of dynamic electricity pricing, in this case households' decisions of whether to switch to a dynamic tariff or pricing structure. They motivate their work on the basis of the importance of better understanding of the determinants of consumers' decisions of whether to adopt dynamic pricing. They argue that, given both the high cost of empirical research on consumer adoption and the special capabilities of ABM for studying it, ABM is not merely an acceptable method but a preferred method for analyzing the adoption problem.

The paper describes the rationale for, structure of, and computational experiments with an ABM of customer adoption of dynamic pricing. The model is based on the idea that decisions are a function of the joint effects of influences on consumers' attitudes and of the degree of indifference with which they regard the desirability of switching. The model is characterized as "hypothetical yet plausible" (p. 172). The paper contains an extensive background review of sociological and other research on the effects of consumer attitudes and social influences on behavior (including energy- and environment-related behavior); electricity use and ways of altering it; and consumer responses to dynamic electricity pricing.

The model itself takes the form of a "social system" represented by a square grid in which each cell corresponds to an agent, called a spin-son (or "spinson") because of its basis in "spin models" in statistical physics. Each spinson has a value of either –1 (a preference for a flat tariff) or +1 (a preference for the dynamic). Initially, each spinson is assumed to have a flat tariff, but it may switch, and the dynamics of the switching behavior in the population is the focus of the simulation.

Notwithstanding the long background discussion mentioned above, the authors acknowledge that there is not "one established theory on how exactly a decision is made" (p. 168). Instead, referring informally to research on the relationship of attitudes to behavior and the importance of attitude "stability," they assert that "consumers have to be really convinced before making a decision" (p. 167). Spinsons' attitudes to the dynamic tariff are either "positive" (favorable) or "negative" (unfavorable). Starting in state –1 (flat tariff), a positive opinion for a sufficient length of time will result in a switch; starting in state +1 (dynamic tariff), a negative opinion for a sufficient length of time will result in a switch back to the flat tariff.

The authors posit that a spinson's opinion at any time is a function of three factors: The level of "indifference," the degree of "conformity," and "product features." A parameter $p \in [0,1]$ represents the level of indifference. A brief description seems to say that conformity is represented by defining a topology in which a spinson has "neighbors" and its opinion is affected by theirs. A parameter $h \in [0,1]$ represents the degree of "influence." A parameter $\tau > 0$ represents the time a spinson takes to make a decision. The numerical values assigned to p and to h are the same for all spinsons and represent "average [societal] values."

Monte Carlo simulations are run for three values of τ and four values of both p and h. The computational analysis entails the emergence of "clusters" of spinsons in the topology generated by the model as the number of simulations increases and the relationships between indifference and decisions.

The authors make several recommendations on the basis of their results, including the need to "communicate . . . the potential benefits of adoption," "provide clear and full information to customers," and "provid[e] enough incentives that will overcome the cost and discomfort when switching to a new tariff" (p. 173).

Because this work both is presented as an alternative to empirical research and claims to draw on and be supported by previous work in a number of disciplines, assessing it should begin with a close examination of the relationship between ABM and analysis and of the research cited to support it.

First, the basic rationale for dynamic pricing is that it will improve the economic efficiency of the electric power system by aligning retail electricity prices with marginal costs of generation, which vary over a 24-hour period. That is in contrast with "green" electricity pricing, which refers to tariffs that electricity consumers voluntarily pay to be supplied by renewable or low-carbon generation sources. A secondary rationale is the potential usefulness of dynamic pricing in facilitating the operation of the power system with increased renewable generation, but this is a complex technical topic that consumers might not easily grasp or relate to their own electricity use. Indeed, in one of the paper's few cited studies on consumer views on dynamic (as opposed to green) pricing, only half the respondents indicated that they were motivated by environmental issues (Dütschke and Paetz, 2013). In another of the cited studies, only 10 percent of respondents were reported as saying that "environmental benefits played a key role" in their adoption decision (Paetz et al., 2012, p. 33). Yet another cited study reported that "environmental reasons were not motivating" customers of an American utility who were surveyed on their views on adopting dynamic pricing (Star et al., 2010, p. 265). Thus, not only is the work cited by the authors on green-pricing adoption of questionable relevance, but the cited work on dynamic pricing is contrary to the suggestion that "green" values or priorities among consumers are relevant to dynamic-pricing adoption.

Second, much of the other work that the authors cite focuses on the adoption of energy-efficient household equipment and on reducing energy use itself. However, dynamic pricing is not itself an "energy-saving" mechanism: Higher prices during hours in which electricity demand is relatively high are intended to reduce demand during these hours, but this may occur through consumers' shifting consumption to off-peak periods rather than reducing their total electricity consumption during a 1-day cycle.[22] That contrasts, for example, with a carbon-emission tax, which—if passed along to electricity consumers in the form of higher retail prices—would be expected to reduce overall energy consumption. Here again, one of the cited studies (Paetz et al., 2012) found that consumers emphasized *monetary* benefits as a motivation to adopt, with the understanding that the benefits would be achieved at least initially through changing behavior—that is, shifting their consumption—not necessarily through reducing consumption by purchasing energy-efficient technology or otherwise.

Similarly, as evidence of the importance of social influence on energy behavior, the authors cite recent work documenting that providing a household with comparative information on energy use—relative to neighbors, the community, and so on—can be effective in reducing use (Allcott, 2011; Ayres et al., 2013). But the likely normative underpinning of that effect—that energy saving is a worthy social and environmental goal—is not obviously present in the case of dynamic pricing, in which individual or household costs and benefits are primary factors. Moreover, Kowalska-Pyzalska et al. (2014, p. 169) themselves acknowledge that

> it is highly unlikely that people actually know the decisions of their neighbors when it comes to electricity tariffs. . . . [Neighbors' electricity bills] are "invisible" (we do not have access to [them]) and we rarely speak about tariffs with our neighbors. Moreover, we may expect people to lie about their energy conservation behaviors.

[22]The extent to which dynamic pricing would lower overall energy use depends on, for example, the local climate and the presence or absence of air conditioning in dwellings.

Turning to the modeling results, the researchers (Kowalska-Pyzalska et al., 2014, pp. 172–173) characterize their analysis as providing

> a hypothetical, yet plausible explanation of why there is such a big discrepancy between consumer opinions, as measured by market surveys, and the actual participation rate in pilot programs and the adoption of dynamic tariffs. . . . Due to a high indifference level in today's retail electricity markets, customer opinions are very unstable and change frequently. This may hamper the adoption of dynamic tariffs.

One might first point out that, reflecting the discussion above, this opinion–participation discrepancy has been reported in the green-pricing, not dynamic-pricing, literature (e.g., Ozaki, 2011). Beyond that, the statement is at best difficult to interpret. That consumer opinions are "unstable" and that this affects adoption is mentioned previously in the paper as a finding of previous research (although neither of the papers cited mentions "opinion instability"). Here, however, the authors seem to be discussing their computational results. Moreover, the claimed causal link to "indifference" also seems to be intended as either a description or an explanation of the model results, inasmuch as "indifference" is one of the input factors in the model. Despite the connotation that the modeling has yielded an *empirically meaningful* finding, on the contrary "indifference" is here only a parameter label, not a defined or measured psychological state of some sort.

Finally, beyond the examples noted above, little of the work cited by the authors to support their modeling and analysis deals with dynamic-pricing adoption specifically (and none of it contributes quantitatively—for example, to parameterizing the ABM).[23] However, although the literature on this topic that is not cited by the authors is relatively small, it is not nonexistent. Moreover, it contains findings that are highly relevant to the ABM analysis in their study. For example, Baladi et al. (1998) found that volunteers for dynamic pricing were distinguished from nonvolunteers by their understanding of their own electricity use patterns and their belief in their ability to respond effectively to the new rates. Similarly, Ericson (2011) found that adoption was more likely among consumers who have home energy-management systems that would help them to adapt effectively to the dynamic rates. Lineweber (2011) reported survey results in which consumers rated "having more control over electricity use" and "reducing bills by avoiding peak use" (p. 95) as the most important motivations for adopting dynamic pricing. A recent consumer behavior study reported findings that included the importance of "opt-out" versus "opt-in" program designs and that the adopters' primary motivation is financial,

[23] Among nearly 80 papers in the reference list, this writer counted three on dynamic-pricing adoption specifically (not including the authors' cited previous work).

that is, reducing their own electricity expenditures (Todd et al., 2013). In part on the basis of such work, the types of actions and policies that the agent-based modelers recommend to increase adoption have in fact long been recognized and promoted, including customer-knowledge hurdles, the importance of effective information provision, and the need for more effective marketing and recruitment approaches (e.g., Barbose et al., 2004). That is, their modeling provides no new or useful practical information on encouraging adoption of dynamic pricing.

Discussion

The two studies just reviewed illustrate the epistemological issues discussed previously but also provide cautionary tales for ABM and its potential use in policy making. The extensive detail in the Ehlen et al. (2007) model is claimed to yield "high fidelity," but no argument or evidence is provided to support the claim. Moreover, data used to populate and parameterize parts of the model are average values and are far from the level of resolution that would support the model's level of mathematical detail. Those are examples of the "complexity and validity" problem. (Also, as noted, sources of several key parameter values are not given at all.)

By using only several or single values of parameters that are identified as having values that could lie anywhere within specified ranges, the Kowalska-Pyzalska model embodies unaddressed Knightian uncertainty. But other aspects of the model and the study are more troublesome. The researchers explicitly claim that their modeling can substitute for empirical research. However, notwithstanding extensive citations of other work that is claimed to support the model's assumptions and structure, close examination reveals that little of that work is directly relevant to or provides evidence for the actual ABM. At the same time, existing work on dynamic pricing that bears directly on the model and the analysis is not cited, and some of it contains evidence counter to the assumptions of the model. Findings and policy recommendations that are generated by the ABM and presented implicitly by the modelers as "new" are on the contrary well known and discussed in the literature. All in all, this study not only provides stark evidence *against* the use of ABM in lieu of empirical research but illustrates the risks of using ABM for policy analysis.

FUNDAMENTAL MODEL UNCERTAINTY AND ROBUSTNESS

The introduction to this Appendix noted nuances associated with the term *validation* in computational energy modeling. An emergent view regarding modeling for regulatory purposes in general (not just energy-focused) is that traditional validation—in particular, as practiced in the physical sciences—is impossible for models of the complex natural systems that are typically the objects of regulation. In this view, "evaluation" is the appropriate alternative (NRC, 2007, p. 3):

> Model evaluation is the process of deciding whether and when a model is suitable for its intended purpose. This process is not a strict validation or verification procedure but is one that builds confidence in model applications and increases the understanding of model strengths and limitations.

This description can be interpreted as an attempt to articulate a decision-theoretic approach to model evaluation in this domain, which has the potential to address the fundamental issues described above. It is a promising direction but has not been systematically operationalized in regulatory modeling (Sanstad, forthcoming). At the most basic level, for this concept of evaluation to be meaningful, *confidence, understanding, strength,* and *limitation* must be defined—a difficult task that for the most part remains to be carried out.[24]

Nevertheless, a decision-oriented approach to modeling is a promising direction. A general framing is that a decision maker either has a model of a system that she or he does not fully trust or is faced with multiple models and does not know which is the correct or otherwise most credible or appropriate one. But the decision maker seeks to use the models, for example, to design policies that affect the system. Macroeconomists have developed and applied several frameworks for analyzing this general problem of "fundamental model uncertainty" (Brock et al., 2003, 2007; Hansen and Sargent, 2008). They differ technically, but they share a focus on "robustness analysis": the study of how decision makers can make choices that will yield outcomes that are acceptable even if not necessarily optimal, given that the model used to predict the outcomes cannot be verified as being "true." Several researchers have adapted those ideas to energy modeling (e.g., Cai and Sanstad, 2014; Loulou and Kanudia, 1999).

A complementary approach to robustness analysis—in effect addressing the Knightian uncertainty problem discussed earlier in this paper—has been developed and implemented by Lempert and colleagues (e.g., Lempert, Popper, and Bankes, 2003; Lempert, Bryant, and Bankes, 2008). (See also Dalal et al., 2013.) In this approach, the space of model solutions generated by using an entire set of equally plausible inputs is computationally generated and explored by using machine learning, visualization, and other

[24]It is important to point out that in contrast with the situation in energy modeling and social science–based and policy modeling more generally, validation, verification, and uncertainty quantification in computational modeling in the physical and engineering sciences have become active and productive topics of basic and applied research. Oberkampf and Roy (2010) is a recent and authoritative source on the subject.

techniques. The output space is analyzed to find "regions" of similar or equivalent policy-relevant outcomes with respect to criteria that reflect a decision maker's preferences and that can be used to identify decisions (represented by input parameter values) that yield outcomes that are robust to the underlying (Knightian) uncertainty.

The idea of fundamental model uncertainty is in principle useful for assessing ABMs. As discussed above, there is a generic problem in ABM of specifying "plausible" agent decision rules when (possibly many) others might be equally justified. A model-uncertainty perspective could facilitate the explicit analysis of this problem and provide firmer grounding for, and increased credibility of, ABMs. This point is taken up in the next section.

LESSONS FOR AGENT-BASED MODELING AND RECOMMENDATIONS FOR REGULATORY APPLICATIONS

In its present stage of development, ABM is a heterogeneous field in many respects, including the level of attention paid among its sub-disciplines to empirical foundations, validation, and uncertainty quantification. As discussed by, for example, Fagiolo et al. (2007) and Windrum et al. (2007), work on such issues is active in some quarters of the ABM community. Bianchi et al. (2007, 2008) describe a rigorous validation and empirically grounded calibration analysis of the agent-based Complex Adaptive Trivial System economic model of financial and capital productivity dynamics in a population of firms. The wider use in ABM of the types of methods that they describe could facilitate rigorous quantification and analysis of, for example, the "bias–variance" trade-off mentioned above and the relationship between complexity and uncertainty in ABMs. However, as the dynamic-pricing examples discussed above demonstrate, concern with such issues is not universal among ABM modelers. (Those examples also demonstrate that ABMs cannot be excused from scrutiny only because they contain a high level of detail or generate "interesting" results.[25])

Lessons

Three key lessons can be drawn from the discussion of energy modeling in this paper. We first note that much of ABM is clearly "calibrationist": it focuses on the computational implications of sets of assumptions rather than on the validity of the assumptions themselves. Unlike the case of standard economic modeling, however, there is for the most part no general, underlying theoretical framework to guide the specification of

[25] See also McNamara et al. (2011) for a critical review of social and behavior modeling, including ABM, specifically for defense applications.

the assumptions. So rather than being—in the phrase of Dawkins et al., 2001—"numerical implementations of theoretical structures" (p. 3672), ABMs in many cases are numerical implementations of *particular assumptions*, which are justified by being plausible or interesting. Thus, in contrast with calibrated economic models, ABMs with some exceptions cannot draw on a large body of knowledge about and insight into a theoretical core of ideas regarding the behavior of agents and the aggregate consequences of their interactions. Those considerations raise the question of whether the computational findings of any given ABM are just idiosyncratic.

Second, although, as discussed, energy modeling is mostly organized around temporal scenario analysis whereas ABM is generally atemporal, there is a basic formal equivalence between the two methods: the logic and structure of energy-scenario modeling are quite similar to those of ABM simulation. The ABM household-electricity examples illustrate that point: The simulations are essentially scenarios of agents' collective behavior. Just as energy modelers disclaim that scenarios are actually predictions, Kowalska-Pyzalska (2014), for example, characterized their model results as "hypothetical but plausible." This again reflects the calibrationist philosophy of focusing on the consequences of underlying assumptions rather than on the assumptions themselves. Thus, like energy modeling, ABM involves a high degree of Knightian uncertainty by virtue of typically not exploring entire plausible input parameter spaces. The characteristic computational intensity of ABM reflects the computational demands of simulating the behavior of agent populations given specific values of key input parameters.

Third, the identification of a high level of model detail or "resolution" with validity or verisimilitude is a hallmark of ABM. However, as in energy modeling, in the absence of explicit criteria for validity, there are no particular theoretical or empirical grounds for this belief that high level of detail entails high level of validity; this is illustrated by the Ehlen et al. (2007) model. Such criteria are generically lacking in ABM. As a consequence, ABM is particularly vulnerable to the illusion-of-precision problem, that is, incorporating great detail that may be only for detail's sake.

Recommendations

As in energy modeling, there is no well-developed, systematic, general method for validating or evaluating ABMs, nor do there appear in general to be topic-specific techniques that might be used to assess ABM for regulatory applications, including applications for tobacco. But several basic questions should be addressed in considering such applications.

First, in a given model, where do the rules governing agents' actions come from? What other equally plausible rules might be used? Why are the ones chosen better justified or more informative than others?

Second, what are the sources of numerical parameters and other model inputs? It has been claimed that an advantage of ABM is that it enables meaningful simulations to be conducted with limited or even no data. But any computational model requires the use of actual numerical values for inputs, which should be well documented and fully justified. The electricity (dynamic-pricing) examples discussed in this paper illustrate the problem.

In this regard, as noted previously, sensitivity analysis can provide useful information about the behavior of a model and can help in interpreting its quantitative output. But it is not a model validation technique and is not a means of justifying particular values for model inputs. A finding that model outputs are relatively insensitive to particular input choices, for example, is instructive but should not be taken as evidence about model validity one way or the other.

The two previous points are illustrated in an ABM of addictive behaviors developed by Moore et al. (in press, Table 1). A "discussion of choice for value" is presented for eight model parameters. For all but one, the justification is based partially or wholly on results of sensitivity analysis.

These issues related to ABM parameter choice and sensitivity analysis and the problem of Knightian uncertainty in ABM indicate the potential usefulness in ABM of the scenario discovery and robustness approach of Lempert and colleagues (e.g., Lempert, Popper, and Bankes, 2003; Lempert, Bryant, and Bankes, 2008). In the Kowalska-Pyzalska and Moore et al. models, each key parameter is associated with a plausible range, but the simulations are based primarily on the use of a single default value or a small number of default values chosen from those ranges. The Lempert et al. methodology of full analysis of model simulations across such ranges would potentially be effective in addressing the parameter choice, sensitivity, and Knightian-uncertainty problems in ABM.

The next question is, what is the intended use of the model output? If it is to provide qualitative insight, careful consideration should be given to whether the insight is into the empirical phenomenon of interest rather than only into the workings of the model itself, including its particular structure and assumptions. That is especially important if the model will be used in establishing a quantitative regulatory rule or criterion. In this case, however, an explicit accounting should be made of whether and how the *quantitative* validity of the model has been demonstrated sufficiently to support this application.

Finally, what are the potential consequences for the regulatory process and its outcomes if the model is wrong? Notwithstanding the manifold difficulties involved in model assessment, validation, or evaluation, this question can and should be addressed by regulators by drawing on their domain-specific knowledge—and common sense.

An example of what can happen if a computational model is wrong is provided by a type of energy modeling not discussed in this paper: real-time and 1-day-ahead electricity modeling that is used in power-system management. In such operational (as opposed to long-range, policy-focused) applications, modeling errors can result in power blackouts.[26] In tobacco regulation, for example, what might be the public health consequences of a modeling error?

CONCLUSION

ABM is an expanding field of social science research and is increasingly considered for use in applied policy analysis and regulation. Thus, it is important for agent-based modelers and their potential constituents to address epistemological issues, such as defining and evaluating model validity, quantifying model uncertainty, and understanding how ABMs should be assessed for use in decision making.

Such issues have a long history in computational energy modeling. This paper is based on the premise that, notwithstanding the differences between the two fields of ABM and computational energy modeling, energy modeling can provide valuable insight and guidance for ABM. It reviewed energy modeling methods, applications, and epistemology, particularly the issues of calibration and its relationship to Knightian uncertainty, the relationship between complexity and model validity, and questions that arise when computational modeling is used for public decision making. It also analyzed several examples of agent-based energy modeling. This review and analysis were then used to draw lessons on validation and uncertainty quantification in ABM and as a basis for recommending guidelines for assessing ABM for regulatory use. It is hoped that the discussion in this paper will contribute to advancing rigorous and policy-relevant assessment of ABM.

[26] Here *error* is used not in a statistical sense but rather to refer to, for example, model misspecification.

Appendix C Annex

A BRIEF HISTORY OF ENERGY-MODEL ASSESSMENT

Through the 1960s, much of "energy modeling" was econometric or statistical regression analysis and short-term forecasting, often applied to energy demand—for example, by electric utility planners. Computational energy modeling became predominant during the 1970s, coinciding with the emergence of energy issues as major public-policy priorities. Environmental effects of energy production and use had been brought to public attention by energy industry experts, scientists, and activists, and the first "oil crisis" sparked by conflict in the Middle East had resulted in extreme concern regarding fuel supplies. "Energy independence" attained high policy and political priority, and one of the first major examples of computational energy modeling during that era was a linear program created for Project Independence, which was initiated by President Nixon in 1973 (Hogan, 1975).

Coinciding with both the continuing attention to energy issues and the policy and regulatory philosophy and priorities of the Carter administration, energy modeling quickly broadened and expanded in number of models, practitioners, and specific applications. Technical work was accompanied by serious, sustained attention to methodological issues, particularly validation, evaluation, and uncertainty quantification. The following are examples. A 1978 bibliography on validation in social science–based and policy modeling, including energy modeling, contained more than 700 entries (Gruhl and Gruhl, 1978). The Energy Modeling Forum (EMF) was created at Stanford University in 1976 "to provide a structural framework within which energy experts, analysts, and policymakers could meet to improve their understanding of critical energy problems" (EMF, 1988, p. i).[27] Concurrently, a model-analysis project was established at the Massachusetts Institute of Technology to conduct rigorous third-party validation of energy-model codes (MIT, 1978). Summary reports of studies organized by the National Bureau of Standards around 1980 reveal a diverse group of scholars and practitioners grappling with fundamental and challenging issues in energy-model evaluation, validation, and uncertainty quantification (e.g., NBS, 1980).

By the end of the 1980s, however, academic and government interest in and resources devoted to energy-model evaluation and related topics had attenuated, following the end of that era of "energy crises" and a politi-

[27]The EMF pioneered systematic, quantitative model comparison and continues as the world's leading center for such activities, playing an important and influential role in energy modeling and policy.

cal shift away from energy regulation and policy at the national level. But energy modeling itself continued in the government, national laboratories, and universities, and in the succeeding decades it steadily expanded and increased in importance for policy applications. As noted in this paper, it has come to dominate quantitative energy-policy analysis, including the subject of carbon-emission abatement.

A 2008 study by the Congressional Research Service (CRS) is a notable exception to the contemporary paucity of attention to energy model validation. It was prepared as background on energy modeling for U.S. Senate hearings on prospective GHG-reduction legislation (Parker and Yacobucci, 2008). At the Senate's request, six energy-modeling studies of potential long-run—to the year 2050—costs and benefits of the proposed policies were conducted. CRS observed that

> It is difficult (and some would say unwise) to project costs up to the year 2030, much less beyond. The already tenuous assumption that current regulatory standards will remain constant becomes more unrealistic, and other unforeseen events (such as technological breakthroughs) loom as critical issues which cannot be modeled. *Long-term cost projections are at best speculative, and should be viewed with attentive skepticism* [emphasis in original]. (Parker and Yacobucci, 2008, p. 73)

CRS does allow that "despite models' inability to predict the future, cases examined here do provide insights on the costs and benefits [of the proposed legislation]" (p. i). The report also quoted a 1990 CRS study of the potential effects of the prospective "cap and trade" program of sulfur dioxide emission from electric power plants:

> [Long-run] cost projections . . . are more a function of each model's assumptions and structure than they are of the details of proposed legislation. **Projections this far into the future are based more on philosophy than analysis** [emphasis in original]. (Parker and Yacobucci, 2008, p. 10)

That is a rare but frank and accurate assessment of the limitations of energy modeling for long-run policy analysis, its primary application. It is especially noteworthy that these conclusions were reached by analysts who were reporting to Congress, a primary constituency for these models' outputs.[28]

[28] In keeping with the federal budgeting process, Congress has been a primary funder of energy modeling, through executive agency budgets and research grants to universities and national laboratories.

REFERENCES

Allcott, H. 2011. Social norms and energy conservation. *Journal of Public Economics* 95(9–10):1082–1095.

Ayres, I., S. Raseman, and A. Shih. 2013. Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. *Journal of Law, Economics, and Organization* 29(5):992–1022.

Baladi, S. M., J. A. Herriges, and T. J. Sweeney. 1998. Residential response to voluntary time-of-use electricity rates. *Resource and Energy Economics* 20(3):225–244.

Barbose, G., C. Goldman, and B. Neenan. 2004. *A survey of utility experience with real time pricing*. Berkeley, CA: Lawrence Berkeley National Laboratory.

Barton, D. C., E. D. Eidson, D. A. Schoenwald, K. L. Stamber, and R. K. Reinert. 2000. *Aspen-EE: An agent-based model of infrastructure interdependency.* Albuquerque, NM: Sandia National Laboratories.

Bianchi, C., P. Cirillo, M. Gallegati, and P. A. Vagliasindi. 2007. Validating and calibrating agent-based models: A case study. *Computational Economics* 30(3):245–264.

———. 2008. Validation in agent-based models: An investigation on the CATS model. *Journal of Economic Behavior & Organization* 67(3–4):947–964.

Borenstein, S. 2013. Effective and equitable adoption of opt-in residential dynamic electricity pricing. *Review of Industrial Organization* 42(2):127–160.

Brock, W. A., S. N. Durlauf, and K. D. West. 2003. Policy evaluation in uncertain economic environments. *Brookings Papers on Economic Activity* 1:235–322.

———. 2007. Model uncertainty and policy evaluation: Some theory and empirics. *Journal of Econometrics* 136(2):629–664.

Burnham, K. P., and D. R. Anderson. 2002. *Model selection and multimodel inference: A practical information-theoretic approach*. 2nd ed. New York: Springer-Verlag.

Cai, Y., and A. Sanstad. 2014. Model uncertainty and energy technology policy: The example of induced technical change. RDCEP Working Paper No. 14-01.

Clarke, L., J. Edmonds, H. Jacoby, H. Pitcher, J. Reilly, and R. Richels. 2007. *Scenarios of greenhouse gas emissions and atmospheric concentrations: Sub-report 2.1a of synthesis and assessment product*. Washington, DC: Climate Change Science Program and the Subcommittee on Global Change Research. Department of Energy, Office of Biological & Environmental Research.

Dalal, S., B. Han, R. Lempert, A. Jaycocks, and A. Hackbarth. 2013. Improving scenario discovery using orthogonal rotations. *Environmental Modelling & Software* 48:49–64.

Dawkins, C., T. N. Srinivasan, and J. Whalley. 2001. Calibration. In *Handbook of econometrics*. Vol. 5, edited by J. J. Heckman and E. Leamer. Amsterdam, Netherlands: Elsevier. Pp. 3653–3703.

Dütschke, E., and A. G. Paetz. 2013. Dynamic electricity pricing—which programs do consumers prefer? *Energy Policy* 59:226–234.

Ehlen, M. A., A. J. Scholand, and K. L. Stamber. 2007. The effects of residential real-time pricing contracts on Transco loads, pricing, and profitability: Simulations using the N-ABLE™ agent-based model. *Energy Economics* 29(2):211–227.

EIA (Energy Information Administration). 2009. *The national energy modeling system: An overview.* Washington, DC: Energy Information Administration, Office of Integrated Analysis and Forecasting, U.S. Department of Energy.

———. 2012. *Analysis of the Clean Energy Standard Act of 2012*. Washington, DC: Energy Information Administration, U.S. Department of Energy.

———. 2014. *Annual energy outlook 2014 with projections to 2040*. Washington, DC: Energy Information Administration, U.S. Department of Energy.

EMF (Energy Modeling Forum). 1988. *North American natural gas markets: EMF report 9*. Palo Alto, CA: Energy Modeling Forum, Stanford University.

EPA (U.S. Environmental Protection Agency). 2009. *EPA analysis of the American Clean Energy and Security Act of 2009: H.R. 2454 in the 111th Congress*. Washington, DC: U.S. Environmental Protection Agency, Office of Atmospheric Programs.

———. 2010. *Documentation for EPA base case v.4.10 using the integrated planning model*. Washington, DC: U.S. Environmental Protection Agency, Clean Air Markets Division.

Ericson, T. 2011. Households' self-selection of dynamic electricity tariffs. *Applied Energy* 88(7):2541–2547.

Fagiolo, G., A. Moneta, and P. Windrum. 2007. A critical guide to empirical validation of agent-based models in economics: Methodologies, procedures, and open problems. *Computational Economics* 30(3):195–226.

Gerst, M. D., P. Wang, A. Roventini, G. Fagiolo, G. Dosi, R. B. Howarth, and M. E. Borsuk. 2013. Agent-based modeling of climate policy: An introduction to the engage multi-level model framework. *Environmental Modelling & Software* 44:62–75.

Goettle, R. J., M. S. Ho, D. W. Jorgenson, D. T. Slesnick, and P. J. Wilcoxen. 2007. IGEM, an Inter-temporal General Equilibrium Model of the U.S. economy with emphasis on growth, energy and the environment. Prepared for the U.S. Environmental Protection Agency, Office of Atmospheric Programs, Climate Change Division.

Goldstein, G. A., L. A. Goudarzi, P. Delaquil, and E. Wright. 2008. *NE-12 MARKAL final report structure, data and calibration: Final deliverable to Northeast States Center for a Clean Air Future under IRG contract*. New York: International Resources Group.

Gruhl, J., and N. Gruhl. 1978. Methods and examples of model validation: An annotated bibliography. Cambridge, MA: MIT Energy Laboratory Working Paper MIT-EL 78–022 WP.

Hansen, L. P. and T. J. Sargent. 2008. *Robustness*. Princeton, NJ: Princeton University Press.

Hogan, W. W. 1975. Energy policy models for Project Independence. *Computers & Operations Research* 2(3–4):251–271.

Intergovernmental Panel on Climate Change. 2000. *Special report: Emissions scenarios—Summary for policymakers*. Cambridge, UK: Cambridge University Press.

Knight, F. 1921. *Risk, uncertainty, and profit*. Chicago, IL: University of Chicago Press.

Koritarov, V. S. 2004. Real-world market representation with agents: Modeling the electricity market as a complex adaptive system with an agent-based approach. *IEEE Power & Energy Magazine* 2(4):39–46.

Kowalska-Pyzalska, A., K. Maciejowska, K. Suszczyński, K. Sznajd-Weron, and R. Weron. 2014. Turning green: Agent-based modeling of the adoption of dynamic electricity tariffs. *Energy Policy* 72:164–174.

Lashof, D. A., S. Yeh, D. Doniger, S. Carter, and L. Johnson. 2013. *Closing the power plant carbon pollution loophole: Smart ways the Clean Air Act can clean up America's biggest climate polluters*. New York: National Resources Defense Council.

Lempert, R. J., S. W. Popper, and S. C. Bankes. 2003. *Shaping the next one hundred years: New methods for quantitative, long-term policy analysis*. Santa Monica, CA: RAND Corporation.

Lempert, R. J., B. P. Bryant, and S. C. Bankes. 2008. *Comparing algorithms for scenario discovery*. Santa Monica, CA: RAND Corporation.

Lineweber, D. C. 2011. Understanding residential customer support for—and opposition to—smart grid investments. *The Electricity Journal* 24(8):92–100.

Loulou, R., and A. Kanudia. 1999. Minimax Regret strategies for greenhouse gas abatement: Methodology and application. *Operations Research Letters* 25(5):219–230.

McNamara, L. A., T. G. Trucano, and C. Gieseler. 2011. *Challenges in computational social modeling and simulation for national security decision making*. Fort Belvoir, VA: Defense Threat Reduction Agency, Office of Strategic Research and Dialogues.

MIT (Massachusetts Institute of Technology). 1978. *Independent assessment of energy policy models: Two case studies*. Cambridge, MA: MIT Energy Laboratory, MIT Model Assessment Group.

Moore, T. W., P. D. Finley, B. J. Apelberg, B. Ambrose, N. S. Brodsky, T. J. Brown, C. Husten, R.J. Glass. In press. An opinion-driven behavioral dynamics model for addictive behaviors. *European Physical Journal B*.

NBS (National Bureau of Standards). 1980. *Validation and assessment issues of energy models: Proceedings of a workshop held at the National Bureau of Standards, Gaithersburg, MD*. Washington, DC: National Bureau of Standards, National Engineering Laboratory, U.S. Department of Commerce.

Nordhaus, W. 2008. *A question of balance: Weighing the options on global warming policies*. New Haven, CT: Yale University Press.

NRC (National Research Council). 1992. *The national energy modeling system*. Washington, DC: National Academy Press.

———. 2007. *Models in environmental regulatory decision making*. Washington, DC: The National Academies Press.

Oberkampf, W. L. R. and C. J. Roy. 2010. *Verification and validation in scientific computing*. New York: Cambridge University Press.

Ozaki, R. 2011. Adopting sustainable innovation: What makes consumers sign up to green electricity. *Business Strategy and the Environment* 20(1):1–17.

Paetz, A.-G., E. Dütschke, and W. Fichtner. 2012. Smart homes as a means to sustainable energy consumption: A study of consumer perceptions. *Journal of Consumer Policy* 35(1):23–41.

Parker, L., and B. Yacobucci. 2008. *Climate change: Costs and benefits of S. 2191. CRS report for Congress RL34489*. Washington, DC: Congressional Research Service.

Ross, M. T. 2005. *Documentation of the Applied Dynamic Analysis of the Global Economy (ADAGE) model: Working paper 05-01*. Research Triangle Park, NC: RTI International.

Sanstad, A. Forthcoming. Abating carbon dioxide emissions from electric power generation: Model uncertainty and regulatory epistemology. *Journal of Legal Studies*.

Star, A., M. Isaacson, D. Haeg, L. Kotewa, and C. Energy. 2010. The dynamic pricing mousetrap: Why isn't the world beating down our door? In *ACEEE summer study on energy efficiency in buildings*. Pp. 257–268.

Todd, A., P. Cappers, and C. A. Goldman. 2013. *Residential customer enrollment in time-based rate and enabling technology programs: Smart grid investment grant consumer behavior study analysis*. Berkeley, CA: Lawrence Berkeley National Laboratory.

Weidlich, A., and D. Veit. 2008. A critical survey of agent-based wholesale electricity market models. *Energy Economics* 30(4):1728–1759.

Windrum, P., G. Fagiolo, and A. Moneta. 2007. Empirical validation of agent-based models: Alternatives and prospects. *Journal of Artificial Societies and Social Simulation* 10(2):8.

Appendix D

Committee Meeting Agendas

Wednesday, February 26, 2014
National Academy of Sciences Building
2101 Constitution Avenue
Washington, DC

10:00–10:15 a.m.	**Welcome and Introductions** Robert Wallace, Committee Chair
10:15–11:00 a.m.	**The Charge to the Committee and Discussion** Shawn L. Fultz, Special Projects Coordinator, Center for Tobacco Products, U.S. Food and Drug Administration
11:00 a.m.–12:00 p.m.	**Agent-Based Tobacco Model Developed for the Center for Tobacco Products, Part 1** Nancy Brodsky, Theresa Brown Model Development Team, Complex Adaptive System of Systems (CASoS) Engineering, Sandia National Laboratories
12:00–12:15 p.m.	**Questions**

1:15–2:15 p.m.	**Agent-Based Tobacco Model Developed for the Center for Tobacco Products, Part 2** Nancy Brodsky, Theresa Brown, Patrick (Pat) Finley Model Development Team, Complex Adaptive System of Systems (CASoS) Engineering, Sandia National Laboratories
2:15–2:45 p.m.	**Committee Questions and Discussion** Nancy Brodsky, Theresa Brown, Patrick (Pat) Finley, Thomas Moore, Stephan (Steve) Verzi Model Development Team, Complex Adaptive System of Systems (CASoS) Engineering, Sandia National Laboratories
3:00–4:00 p.m.	**Overview: An Agent-Based Model to Study Market Penetration of Plug-in Hybrid Electric Vehicles and Discussion** Margaret Eppstein, Chair and Associate Professor, Department of Computer Science, University of Vermont
4:00–5:00 p.m.	**Evidence of Social Interactions in Drug Initiation and Discussion** Jonathan P. Caulkins, Professor of Operations Research and Public Policy, Heinz College of Public Policy and Management, Operations Research Department, Carnegie Mellon University
5:00–5:15 p.m.	**Public Comment**
5:15 p.m.	**Closing Comments/Adjourn Open Session** Robert Wallace, Committee Chair

Thursday–Friday, April 17–18, 2014
National Academy of Sciences Building
2101 Constitution Avenue
Washington, DC

Thursday April 17, 2014

9:00–9:15 a.m.	**Welcome and Introductions** Robert Wallace, Committee Chair
9:15–9:45 a.m.	**Dynamics of Social Networks and Influence** Tom Valente, Professor, Department of Preventive Medicine, University of Southern California
9:45–10:15 a.m.	**Exploring Network Effects and Tobacco Use with SIENA** David Schaefer, Associate Professor, School of Human Evolution and Social Change, College of Liberal Arts and Sciences, Arizona State University
10:15–10:45 a.m.	**Discussion** David Schaefer Tom Valente
11:00–11:30 a.m.	**Social and Behavioral Sciences for Tobacco Use** Joseph Cappella, Gerald R. Miller Professor of Communication, Annenberg School for Communication, University of Pennsylvania
11:30–11:45 a.m.	Bonnie L. Halpern-Felsher (via phone), Professor, Department of Pediatrics, Stanford School of Medicine
11:45 a.m.–12:15 p.m.	**Discussion** Joseph Cappella Bonnie L. Halpern-Felsher
12:15–1:15 p.m.	**Lunch**

1:15–2:45 p.m.	**Agent-Based Models: Lessons and Policy Implications from Relevant Models** Ross Hammond, Senior Fellow and Director, Center on Social Dynamics and Policy, Brookings Institution
2:45–3:30 p.m.	**Panel Discussion—Reflections on the Day/ Follow-Up Questions** Joseph Cappella Ross Hammond David Schaefer Tom Valente
3:30–3:45 p.m.	**Public Comment**
3:45 p.m.	**Closing Comments/Adjourn Open Session** Robert Wallace, Committee Chair

Friday, April 18, 2014

8:45–9:00 a.m.	**Welcome and Introductions** Robert Wallace, Committee Chair
9:00–9:40 a.m.	**Overview of Opinion Dynamics Modeling, Strengths/Weaknesses, Implications for Policy and Clarifying Questions from Committee Members** Sidney Redner, Professor, Physics Department, Boston University
9:40–10:00 a.m.	Andreas Flache (via phone), Professor, Department of Sociology, University of Groningen, The Netherlands
10:00–10:35 a.m.	**Decision Making and Risk Perception** Ellen Peters, Professor of Psychology, Director of the Behavioral Decision Making Initiative, Ohio State University
10:35–11:30 a.m.	**Discussion—All Speakers**
11:30–11:45 a.m.	**Public Comment**

11:45 a.m.	**Adjourn Open Session**

June 26, 2014
Keck Center
500 Fifth Street NW
Washington, DC

8:15–8:30 a.m.	**Welcome and Introductions** Robert Wallace, Committee Chair
8:30–9:15 a.m.	**Agent-Based Tobacco Model Developed for the Center for Tobacco Products—An Update** Nancy Brodsky, Thomas Moore Sandia National Laboratories
9:15–9:45 a.m.	**Discussion/Q&A**
9:45–10:15 a.m.	**Models of Infectious Disease Agent Study (MIDAS)—Examples of Agent-Based Modeling** Joshua Epstein, Professor, Emergency Medicine, Departments of Applied Mathematics, Economics, Biostatistics, International Health, Civil Engineering, Environmental Health Sciences, and the Institute for Computational Medicine, Johns Hopkins University; External Professor, Santa Fe Institute
10:15–10:35 a.m.	**Discussion/Q&A**
10:50–11:20 a.m.	**Finding the Unknown Unknowns: Using Dynamic Models to Identify and Reduce Uncertainty in Regulatory Analyses** Scott Spak, Assistant Professor, Urban and Regional Planning and Civil and Environmental Engineering, Public Policy Center, University of Iowa
11:20–11:40 a.m.	**Discussion/Q&A**

11:40 a.m.–12:15 p.m.	**Understanding Fundamental Model Uncertainties and Their Decision-Making Implications: Lessons from Energy Policy** Alan H. Sanstad, Staff Scientist, Sustainable Energy Systems Group, Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory
12:15–12:40 p.m.	**Discussion/Q&A**
12:40–12:45 p.m.	**Closing Comments/Adjourn Open Session** Robert Wallace, Committee Chair

Appendix E

Committee Biographical Sketches

ROBERT WALLACE, M.D., M.Sc. (*Chair*), is a professor of epidemiology at the University of Iowa College of Public Health, a professor of internal medicine at the university's Carver College of Medicine, and the director of the university's center on aging. He is an elected member of the Institute of Medicine (IOM), where he has previously chaired two boards and participated in many consensus committees. He has been a member of the U.S. Preventive Services Task Force and the National Advisory Council on Aging of the National Institutes of Health (NIH). He is a former chair of the epidemiology section of the American Public Health Association. He is the author or co-author of more than 400 peer-reviewed publications and 25 book chapters, and he has edited or co-edited 4 books, including the current edition of *Maxcy–Rosenau–Last Public Health and Preventive Medicine*. Dr. Wallace's research interests concern the causes and prevention of disabling conditions of older persons. He is a co–principal investigator of the Health and Retirement Study, a long-term prospective sample of older Americans exploring health, social, family, and economic policy issues, and a co-investigator of the Women's Health Initiative (WHI), a national study exploring the prevention of important chronic diseases in older women. He has been a collaborator on several international studies regarding the prevention of chronic illness in older persons. Dr. Wallace is currently a member of the advisory board for the National Research Council's (NRC's) Division of Behavioral and Social Sciences and Education.

ELIZABETH BRUCH, Ph.D., is an assistant professor in sociology and complex systems and an affiliate of the Population Studies Center at the

Institute for Social Research. Her expertise includes decision making, choice modeling, and population dynamics. Much of her work blends statistical and agent-based methods to examine the relationship between individuals' decisions about where to live and patterns of residential segregation. She is also working on a project exploring the implications of mate search strategies and mate choice behavior for dating markets, using data from online dating sites. She earned her Ph.D. from the University of California, Los Angeles, in 2006, and was a Robert Wood Johnson Health Policy Scholar from 2006 to 2008. Her article on racial tolerance and race–ethnic segregation, "Neighborhood Choice and Neighborhood Change," won the 2005–2006 Gould Prize; the James S. Coleman Best Article award from the mathematical sociology section of the American Sociology Association (ASA); and the Robert Park Best Article award from the community and urban sociology section of the ASA.

KAREN GLANZ, Ph.D., M.P.H., is the George A. Weiss University Professor in the Perelman School of Medicine and the School of Nursing and the director of the University of Pennsylvania Prevention Research Center and the Center for Health Behavior Research at the University of Pennsylvania. Her research focuses on cancer prevention and control, theories of health behavior, obesity and the built environment, social and health policy, and new health communication technologies. She is currently conducting research on skin cancer prevention, nutrition and chronic disease prevention, compliance with glaucoma medications, and colorectal cancer screening. Her research and community programs in tobacco use prevention have included monitoring and reducing youth access to tobacco, developing a youth advocacy anti-tobacco program, and leading a cluster-randomized trial of a school-based youth activation program for preventing tobacco use. Dr. Glanz and her team are committed to conducting scientific research with promising short- and long-term applications to improved community health, health care, and public health services. She was formerly at Emory University (2004–2009), the University of Hawaii (1993–2004), and Temple University in Philadelphia (1979–1993). Dr. Glanz received her M.P.H. (1977) and Ph.D. (1979) degrees in health behavior and health education from the University of Michigan. She has been recognized with several national awards, and she was the 2007 recipient of the Elizabeth Fries Health Education Award from the James and Sarah Fries Foundation. She is a member of the Task Force on Community Preventive Services, a federally appointed task force that oversees the *Community Guide* evidence reviews. Her scholarly contributions consist of more than 400 journal articles and book chapters. Dr. Glanz is senior editor of *Health Behavior and Health Education: Theory, Research, and Practice* (Jossey-Bass Inc., 1990, 1996, 2002, 2008), a widely used text recently published in its fourth edition. She was recognized in 2006 as a highly cited author by ISIHighlyCited.com, in the top 0.5 percent of authors in her field over a 20-year period. She was elected to the IOM in 2013.

CARLOS ROBERTO JAÉN, M.D., Ph.D., is the Dr. and Mrs. James L. Holly Distinguished Professor and chairman in the Department of Family and Community Medicine and a professor of epidemiology and biostatistics at the University of Texas Health Science Center, San Antonio. Dr. Jaén's research interests focus on improving preventive care for individuals of all ages and preventing complications from chronic diseases such as diabetes, high blood pressure, and heart disease. He is passionate about building and studying high-performance primary care offices. He served on the panels that published smoking cessation guidelines in 1996 and 2000 and was a co-chair of the panel that published an update in May 2008. In 2005 he was appointed to the National Advisory Council to the Agency for Healthcare Research and Quality. He was a co-director of the Center for Research in Family Medicine and Primary Care. The center studied more than 500 mostly independent primary care practices over the past 20 years. Dr. Jaén has his Ph.D. in epidemiology and community health from the State University of New York (SUNY) Buffalo, and his M.D. from the same university. He also has an M.S. in oncology. He received a Generalist Physician Faculty Scholar Award from the Robert Wood Johnson Foundation and a Cancer Control Career Development Award for Primary Care Physicians from the American Cancer Society. He is a practicing family physician and has been selected to the Best Doctors in America yearly since 2002. His interests include building a healthier San Antonio through efforts in community wellness. He was elected to the IOM in 2013 and served on the IOM/NRC Committee on Geographic Adjustment Factors in Medicare Payment.

SCOTT J. LEISCHOW, Ph.D., is a professor of health sciences research at the Mayo Clinic (Mayo Foundation for Medical Education and Research). His expertise and research interests are focused on tobacco treatment and the role of social networks in the dissemination and implementation of evidence-based practices. His interests include optimal approaches for training health care providers on treating tobacco dependence and understanding how complex systems can be understood and optimized to foster improvements in health, particularly in underserved populations in the United States and globally. Currently his work is focused on investigating pharmacological and behavioral treatments for tobacco dependence, including the exploration of dosing, adherence, and the interplay of clinical and public health approaches; researching organizational networks to better understand how knowledge is spread; and implementing practices that can lead to improvements in public health. Prior to his role at the Mayo Clinic, Dr. Leischow was a professor of family and community medicine and the associate director for biobehavioral and social sciences research at the Arizona Cancer Center, University of Arizona. He also served as the chief of the Tobacco Control Research Branch at the National Cancer Institute (NCI). He was seconded from the NCI to the Office of the Secretary of the U.S. Department of Health and Human Services and served as a senior advisor for tobacco policy. Dr. Leischow received his M.A. and his Ph.D. from the University of Maryland, College Park. His research-related memberships include member and advisory board, Arizona Tobacco Revenue, Use, Spending and Tracking Commission; member, American Association for Cancer Research, Tobacco and Cancer Subcommittee; member and past president, Society for Research on Nicotine and Tobacco; and member, Society for Behavioral Medicine. He is the recipient of multiple awards, including the NIH Director's Award (2004), an NIH Individual Merit Award (2003), and the NCI Director's Gold Star Award (2004).

DOUGLAS LUKE, Ph.D., is a professor at Washington University in St. Louis at the George Warren Brown School of Social Work and the director of the Center for Public Health Systems Science. He is a leading researcher in the areas of health behavior, organizational and systems science, public health policy, and tobacco control, and he is also a top biostatistician and social science methodologist who has made significant contributions to the evaluation of public health programs, tobacco control and prevention policy, and the application of new methods to community health interventions. Dr. Luke has expanded the repertoire of statistical methods, such as the use of social network analysis and multilevel models, in the field of public health. He co-directed the doctoral program at the Saint Louis University School of Public Health. He is active in the American Statistical Association, the International Network for Social Network Analysis, and the Society for Community Research and Action. He received his M.A. and Ph.D. in psychology from the University of Illinois. He is a past member of the NIH Center for Scientific Review study section Community Influences on Health Behavior, has published 1 of the 10 most influential methodology articles published in the first 25 years of the *American Journal of Community Psychology*, "Expanding Behavior Setting Theory: Setting Phenotypes in a Mutual Help Organization," and is the recipient of the 1989 Ed Scheiderer Memorial Award, Outstanding Research by a Clinical Psychology Graduate Student, University of Illinois.

CHARLES F. MANSKI, Ph.D., is the Board of Trustees Professor of Economics at Northwestern University, where he specializes in econometrics and social policy. Dr. Manski's research spans econometrics, judgment and decision, and the analysis of social policy. He is the author of *Public Policy* *in an Uncertain World*, *Identification for Prediction and Decision*, and *Identification Problems in the Social Sciences*. He has served as the director of the Institute for Research on Poverty and the editor of the *Journal of Human Resources*. Dr. Manski received his Ph.D. from the Massachusetts Institute of Technology. He is an elected member of the National Academy of Sciences and a fellow of both the Econometric Society and the American Academy of Arts and Sciences. He has served on numerous committees for the NRC and the IOM and currently serves on the Report Review Committee, the Committee on Law and Justice, the Committee for the Context of Military Environments: Social and Organizational Factors, and the Committee on Intercity Passenger Travel: Opportunities and Issues in Short-Haul Markets.

DAVID MENDEZ, Ph.D., is an associate professor in the Department of Health Management and Policy (HMP) at the University of Michigan School of Public Health. Dr. Mendez is the director of the HMP executive master's program. His research focuses in the application of mathematical/computational models for public health policy, particularly in the field of tobacco control. He has conducted research on the impact of tobacco control policies on smoking prevalence and health outcomes. He has also been involved in research to evaluate policies regarding residential radon. He is currently engaged in a study to evaluate the impact of school mandates on the human papillomavirus vaccine uptake using an agent-based model. His other professional affiliations include the Decision Sciences Institute and the Institute for Operations Research and the Management Sciences. Dr. Mendez received his M.S. in applied statistics and an M.S. in operations research/systems science from Michigan State University (1990, 1987). He holds a Ph.D. in management science, also from Michigan State University (1995).

NANCY A. RIGOTTI, M.D., is an academic general internist who is a professor of medicine at Harvard Medical School and the associate chief of the General Medicine Division at Massachusetts General Hospital (MGH). She is an internationally known expert on tobacco cessation treatment and policy. She has pioneered research on interventions to reduce smoking prevalence and the burden of tobacco-related morbidity and mortality. Her clinical research focuses on developing and disseminating interventions for smoking cessation within primary care practices and inpatient settings. She founded and directs the MGH Tobacco Research and Treatment Center, which combines a clinical service providing smoking cessation treatment with a research group that develops and tests smoking treatment interventions for health care settings. She has received numerous grants to support her work, including funding from the National Heart, Lung, and Blood

Institute; the NCI; the American Cancer Society; and the Robert Wood Johnson Foundation. She has published more than 150 original articles about tobacco issues and has written numerous book chapters, reviews, and editorials on these topics. Dr. Rigotti is a past president of the Society of General Internal Medicine and a past president of the Society for Research in Nicotine and Tobacco. From 1985 to 1990 she served as the associate director of the Institute for the Study of Smoking Behavior and Policy at the John F. Kennedy School of Government at Harvard University. Dr. Rigotti was a scientific editor of the 1989 *Surgeon General's Report on Tobacco*, a comprehensive review that provided scientific support for policy making, and she was a deputy editor of *Nicotine & Tobacco Research*, among other editorial responsibilities. Dr. Rigotti has served on numerous national and international committees addressing tobacco issues, and in 2006 she won the James D. Bruce Memorial Award for Distinguished Contributions in Preventive Medicine from the American College of Physicians. Dr. Rigotti received her M.D. from Harvard Medical School in 1978 and completed her residency in primary care internal medicine at MGH.

DAVID SHOHAM, Ph.D., M.S.P.H., is an associate professor of public health sciences at Loyola University Chicago. Dr. Shoham's current research interests focus on social determinants of obesity and kidney disease. He is a co-principal investigator on the Modeling Obesity Through Simulation project, funded by the National Institute of Child Health and Human Development. This project focuses on peer, family, neighborhood, and school influences on childhood obesity using social network analysis and agent-based modeling and simulation. Dr. Shoham is also a co-investigator on the Modeling the Epidemiologic Transition project, funded by the National Institute of Diabetes and Digestive and Kidney Diseases, where he is studying the relationship of occupation and wealth to energy expenditure among African-origin populations in five contexts (Maywood, United States; Ghana; Jamaica; the Seychelles; and South Africa). Dr. Shoham is the admissions director for the M.P.H. program and graduate program director for the epidemiology track, and he is chair of Loyola's Council of Graduate School Programs. He teaches social epidemiology and an introductory course in epidemiologic methods to M.P.H. students. Dr. Shoham completed his Ph.D. in epidemiology at the University of North Carolina at Chapel Hill in 2005, with an emphasis on life-course social epidemiology and kidney disease.

JODY SINDELAR, Ph.D., is a professor of public health and a health economist in the Department of Health Policy and Management at the Yale School of Public Health. In addition, Dr. Sindelar is a research associate at the National Bureau of Economic Research, a research fellow at IZA (Institute for the Study of Labor), and associated faculty at the Institution for Social and Policy Studies at Yale. She was the Bing Visiting Faculty at RAND Corporation in Santa Monica, California, and Washington, DC, and was a founding member and a previous president of the American Society of Health Economists. Dr. Sindelar is an expert on the economics of obesity and of substance abuse, including alcoholism, the use of illicit drugs, and smoking. Tobacco is her current focus. Her studies examine lost productivity, the cost-effectiveness of treatments, social costs, and policy. She has published more than 100 papers and studies on the impacts of substance abuse on productivity, educational attainment, gender differences, and related policy issues in a variety of economics, policy, addiction, health, and medical journals. She has served on numerous editorial reviews and advisory and other boards and committees, and she has presented her research at seminars and conferences both nationally and internationally. Currently, Dr. Sindelar heads multiyear projects funded by the Centers for Medicare & Medicaid Services via the Connecticut Department of Social Services, the National Institute on Aging (NIA), the National Institute of Alcohol Abuse and Alcoholism (NIAAA), and the National Institute on Drug Abuse (NIDA), and she also collaborates on projects funded by the Agency for Healthcare Research and Quality (AHRQ) and the RAND Corporation via the Harvard School of Public Health. She has also been the principal investigator and a collaborator on numerous past research projects funded by the Robert Wood Johnson Foundation, AHRQ, the National Institute of Mental Health, the Yale Center for Clinical Investigation, the U.S. Department of Veterans Affairs, NIA, NIAAA, and NIDA, among others. Dr. Sindelar received her Ph.D. in economics from Stanford University.

MELISSA STIGLER, Ph.D., M.P.H., is an associate professor at the University of Texas Health and Science Center's School of Public Health. She is with the Austin Regional Campus and the Division of Epidemiology, Human Genetics, and Environmental Sciences. She specializes in the epidemiology of youth tobacco use and in designing interventions for schools and communities to help reduce smoking and other types of tobacco use among young people. She is currently the principal investigator for several studies funded by the National Institutes of Health. These include *¡Actívate Ya!*, which is designed to develop, implement, and evaluate a multicomponent tobacco prevention and physical activity promotion intervention among youth in Uruguay. She currently serves as the associate director of the new Tobacco Center of Regulatory Science for Youth and Young Adults in Texas and leads a large surveillance study of tobacco use behaviors and the impact of tobacco marketing on youth with the center. Dr. Stigler served as the associate editor of the 2012 Surgeon General's report on tobacco use among young people and co-authored the chapter on the epidemiology of tobacco use among youth and adults in the United States for the 50th Anniversary Report, published in 2014. Dr. Stigler received her M.P.H. from the University of Minnesota School of Public Health and her Ph.D. in behavioral epidemiology from the same institution. She is the recipient of more than 15 awards related to her research, including the 2012 Inaugural Grant U.S. Health award from the Dell Center for Health Living and the Young Investigator's Award for Outstanding Research from the Health Sciences Center at the University of Texas. She was the 2009 co-chair of the Global Youth Meet on Tobacco Control at the 14th World Conference on Tobacco or Health.

Index

[Page numbers followed by *b*, *f*, or *t* refer to boxed text, figures, or tables, respectively.]

A

D

E

O

P

Q

R

T

U

V

W

Y